

UCLA

UCLA Electronic Theses and Dissertations

Title

Essays in Education and Crime in Colombia

Permalink

<https://escholarship.org/uc/item/3d17r3fk>

Author

Arteaga, Carolina

Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Essays in Education and Crime in Colombia

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Economics

by

Maria Carolina Arteaga

2019

© Copyright by

Maria Carolina Arteaga

2019

ABSTRACT OF THE DISSERTATION

Essays in Education and Crime in Colombia

by

Maria Carolina Arteaga

Doctor of Philosophy in Economics

University of California, Los Angeles, 2019

Professor Adriana Lleras-Muney, Chair

This dissertation contains three essays in applied microeconomics. In the first chapter paper I test whether the return to college education is the result of human capital accumulation or instead reflects the fact that attending college signals higher ability to employers. I exploit a reform at Universidad de los Andes, which in 2006 reduced the amount of coursework required to earn degrees in economics and business by 20% and 14%, respectively, but did not change the quality of incoming or graduating students. The size of the entering class, their average high school exit exam scores, and graduation rates were not affected by the reform, indicating that selection of students into the degrees remained the same. Using administrative data on wages and college attendance, I estimate that wages fell by approximately 16% in economics and 13% in business. These results suggest that human capital plays an important role in the determination of wages and reject a pure signaling model. Surveying employers, I find that the reduction in wages may have resulted from a decline in performance during the recruitment

process, which led students to be placed in lower-quality firms. Using data from the recruitment process for economists at the Central Bank of Colombia, I find that the reform reduced the probability of Los Andes graduates' being hired by 17 percentage points. The second chapter provides evidence that parental incarceration increases children's educational attainment. I collect criminal records for 90,000 low-income parents who have been convicted of a crime in Colombia, and combine it with administrative data on the educational attainment of their children. I exploit exogenous variation in parental incarceration resulting from the random assignment of defendants to judges with different propensities to convict and incarcerate. I find that conditional on conviction, parental incarceration increases education by 0.7 years for children whose parents are on the margin of incarceration. This positive effect is larger for boys, violent crimes, and cases in which the incarcerated parent is the mother. Finally, in the third chapter I derive a new expression that extends the Local Average Treatment Effect concept, to a setting with two sources of unobserved treatment heterogeneity.

The dissertation of Maria Carolina Arteaga is approved.

Maurizio Mazzocco

Sarah Reber

Till M. von Wachter

Adriana Lleras-Muney, Committee Chair

University of California, Los Angeles

2019

I dedicate this dissertation to my parents, my sisters, my niece and nephew.

Contents

1. The Effect of Human Capital on Earnings: Evidence from a Reform at Colombia's Top University	
1.1. Introduction.....	2
1.2. Theoretical framework.....	6
1.3. Institutional background and reform.....	9
1.3.1. Reform.....	9
1.3.2. First stage: Empirical evidence of the reform for economics and business....	10
1.4. Effects of the Reform.....	13
1.4.1. Data.....	13
1.4.2. Preliminary evidence and empirical strategy.....	14
1.4.3. Results.....	15
1.5. Mechanism.....	18
1.6. Robustness Checks.....	20
1.7. Conclusions.....	24
1.8. References.....	26
1.9. List of Figures.....	30
1.10 List of Tables.....	38
2. The Cost of Bad Parents: Evidence from the Effects of Parental Incarceration on Children's Education	
2.1. Introduction.....	50
2.2. Background: The Colombian Court System.....	55
2.3. Data Construction.....	58
2.3.1. Data Sources.....	58
2.3.2. Sample.....	59
2.3.3. Summary Statistics.....	60
2.4. Identification.....	61
2.4.1. A simplified framework.....	63
2.5. Estimation.....	65
2.5.1. Instrument Validity.....	66
2.5.2. Results.....	68
2.5.3. Heterogeneity.....	70
2.5.4. Robustness.....	72
2.6. Mechanisms.....	73
2.6.1. What explains the positive effect?.....	73
2.6.2. How does the environment of the child change?.....	74
2.6.3. Parents at the margin.....	74
2.6.4. External validity and policy implications.....	75
2.7. Conclusions.....	76
2.8. References.....	77
2.9. List of Tables	84
2.10 List of Figures	92

3. Identification with two sources of unobserved treatment heterogeneity	
3.1. Model.....	102
3.2. References	109

List of Figures

1.1 Effect of the reform in degree duration.....	30
1.2 Effect of the reform in in credits studies.....	31
1.3 Effect of the reform on class selection.....	32
1.4 Pre trends and the effect of the reform on wages.....	33
1.5 Treatment effect distribution (Table 3a and 3b).....	34
1.6 Treatment effect distribution (Table 5).....	35
1.7 Treatment effect distribution (Table 6)	36
1.8 Changes in changes estimates	37
2.1 Prosecution and trial stages.....	92
2.2 Incarceration rates.....	93
2.3 Identification.....	94
2.4 Identification under 4 types of judges.....	95
2.5 Judges' fixed effects.....	96
2.6 First stage.....	97
2.7 Scatter plot: Judges' fixed effects.....	98
2.8 Reduced form.....	99
2.5 MTE.....	100
2.6 Model of parenting and incarceration.....	101

List of Tables

1.1 First stage.....	38
1.2 Summary statistics	38
1.3 Baseline results. Effect of the reform on wages	39
1.4 Placebo coefficients.....	41
1.5 Effect of the reform on wages.....	42
1.6 Cap at three years of experience	43
1.7 Effect of the reform on the recruitment process.....	44
1.8 Placebo test 1.....	45
1.9 Placebo test 2.....	46
1.10 Robustness checks	47
1.11 Synthetic control.....	48
2.1 Population by conviction and incarceration.....	84
2.2 Convicted parents by incarceration and gender.....	85
2.3 First stage – Parents.....	86
2.4 Balance test-Trial sample.....	87
2.5 Balance test II-Incarcerated sample.....	88
2.6 Monotonicity test: Norris.....	89
2.7 Monotonicity Test: Frandsen et al.....	89
2.8 OLS Regression.....	90
2.9 Results: Reduced form and IV.....	90
2.10 Heterogeneous effects.....	91
2.11 Changes after incarceration.....	91

Acknowledgments

I am deeply grateful to many people for their support during the completion of this dissertation. I thank my advisors Adriana Lleras-Muney, Rodrigo Pinto, Sarah Reber, Maurizio Mazocco and Till von Wachter for their guidance. I am also thankful for the time of other faculty, especially Leah Boustan, Ricardo Perez-Truglia, Moshe Buchinsky, Denis Chetverikov, Christian Dippel, Paola Giuliano, Martin Hackman, Rosa Matzkin and Manisha Shah. I am grateful to my classmates, not only for comments on my research, but also for insightful academic discussions in all areas of Economics, and for making my PhD a very happy and rich experience. I am especially indebted with Sepehr Ekbatani, Lucia Yanguas, Richard Domurat, Stefano Fiorin, Alex Fon, Keyoung Lee, Rustin Partow, Ruyao Shi and Andreas Gulyas. As a long-time resident of UCLA's California Center for Population Research (CCPR), I thank CCPR for providing invaluable resources and a vibrant research environment; I am especially grateful for the time and support of Ana Ramirez. Finally, I thank my mother, father, sisters, and friends, for their support throughout the entire doctoral program.

This research was made possible by the following awards: Lauchin Currie Scholarship for Graduate Studies - Colombian Central Bank, Scholarship for Doctoral Studies from Colciencias, CCPR Treiman Fellowship, Marcia & Herbert Howard and Jacqueline & George Mefferd Fellowship, the Latin American Institute fellowship and the Dissertation Year Fellowship through the UCLA Graduate Division.

VITA

Maria Carolina Arteaga

EDUCATION

M.A. Economics

University of California, Los Angeles, 2015

M.A. Economics

Universidad de Los Andes, Bogotá, 2010

B.A. Economics

Universidad de Los Andes, Bogotá, 2008

PUBLICATIONS

“The Effect of Human Capital on Earnings: Evidence from a Reform at Colombia’s Top University” (2018). *Journal of Public Economics*. Vol 157, January. pp 212-225.

"El comportamiento del tipo de cambio real en Colombia: ¿explicado por sus fundamentales?" *Ensayos sobre Política Económica*, (72), December (2013) –with Jair Ojeda Joya & Joan Granados.

“Human Capital Externalities and Growth,” *Ensayos sobre Política Económica*, Bogota (66), Banco de la Republica, December 2011.

Books

“*The Healthcare System*” with Mauricio Santa Maria, Juan Gonzalo Zapata and Carlos Felipe Reyes, Fedesarrollo (in Spanish), Bogota, 2009.

PROFESSIONAL EXPERIENCE

2009 –2013 Central Bank of Colombia Bogotá, Colombia

Economist – Inflation Department

2008- 2009 FEDESARROLLO Bogotá, Colombia

Research Assistant

Chapter 1

The Effect of Human Capital on Earnings: Evidence from a Reform at Colombia's Top University

Carolina Arteaga¹, UCLA

In this paper I test whether the return to college education is the result of human capital accumulation or instead reflects the fact that attending college signals higher ability to employers. I exploit a reform at Universidad de los Andes, which in 2006 reduced the amount of coursework required to earn degrees in economics and business by 20% and 14%, respectively, but did not change the quality of incoming or graduating students. The size of the entering class, their average high school exit exam scores, and graduation rates were not affected by the reform, indicating that selection of students into

¹ Department of Economics, UCLA. Contact information: caroartc@ucla.edu. I am grateful to the Colombian Ministry of Education, the Central Bank of Colombia, and the Economics Department at Universidad de los Andes for providing the data for this study. I would like to thank Magne Mogstad and two anonymous referees for their excellent comments. I am extremely grateful to Adriana Lleras-Muney for her encouragement and suggestions. I also want to thank David Atkin, Leah Boustan, Moshe Buchinsky, Michela Giorcelli, Carlos Medina, Maurizio Mazzocco, Rodrigo Pinto, Sarah Reber, Juan E. Saavedra, Andres Santos, and Till von Wachter for their comments and feedback. I am grateful to my colleagues Pasha Andreyanov, Tiago Caruso, Richard Domurat, Keyoung Lee, Rustin Partow, and Maria Lucia Yanguas for insightful suggestions and discussions. I thank seminar participants at UCLA, SOLE, LACEA, EBE, Universidad de Los Andes and the Central Bank of Colombia for valuable comments.

the degrees remained the same. Using administrative data on wages and college attendance, I estimate that wages fell by approximately 16% in economics and 13% in business. These results suggest that human capital plays an important role in the determination of wages and reject a pure signaling model. Surveying employers, I find that the reduction in wages may have resulted from a decline in performance during the recruitment process, which led students to be placed in lower-quality firms. Using data from the recruitment process for economists at the Central Bank of Colombia, I find that the reform reduced the probability of Los Andes graduates' being hired by 17 percentage points.

1.1 Introduction

Education is one of the most important determinants of wages at the individual level. Returns to a year of schooling are estimated to be positive and large in most countries, ranging from 2% to 20% around the world (Montenegro and Patrinos, 2014). Moreover, the earnings premium associated with college has risen substantially in the last decades (Oreopoulos and Petronijevic, 2013). There is less consensus about the mechanisms through which education leads to higher wages. Studies that estimate causal returns to schooling cannot shed light on the sources of such returns (Card, 2001). Two main channels have been proposed in the literature. First, the human capital theory argues that education increases productivity and wages rise as a result (Becker, 1964 and Mincer, 1974). Second, the signaling theory posits that higher wages reflect the correlation between education and unobserved ability.² In both settings, higher-ability workers

² Of course, the two theories are not mutually exclusive.

obtain higher levels of schooling and are paid more, which explains the difficulty in setting the two theories apart.

In this paper, I identify the effect of human capital accumulation on wages, separate from that of signaling, by exploiting a curriculum change at Universidad de los Andes, the top university in Colombia. In 2006, the number of credits required to earn a college degree in economics and business decreased by 20% and 14%, respectively. This was accomplished by dropping 12 required courses in economics and 6 in business, and a reduction in instruction time from 4.5 to 4 years.³ The identification strategy of this paper relies on the fact that the reform did not alter the selection of entering or graduating students. At Los Andes, the admission process is constrained by a limited number of slots and is solely based on scores on the national standardized high school exit exam (the Saber 11). I show that the size of the entering class did not grow, nor did average entrance test scores decrease, and dropout rates did not change with the reduction in the number of classes. Therefore, the reform had no short-run effect on the quality of the entering class after 2006, but it decreased human capital accumulation. The human capital model predicts a decline in wages as a result of the reform, whereas a pure signaling model does not.

To estimate the effect of the reform, I use individual information on wages and educational attainment in a difference-in-differences (DID) framework. I compare wages in the formal sector before and after the reform for economics and business graduates of Los Andes and other top-10 schools in Colombia that did not reform their degrees. I find that after the reform, wages for students from Los Andes decreased by 16% in economics and 13% in business. This

³ In economics, the change in curriculum not only reduced the number of semesters, but also the number of courses per semester. Before the reform, students were expected to take six courses per term; this was changed to five. In business, the number of classes per term remained at five.

suggests that human capital accumulation plays an important role in the determination of wages, and therefore I reject a model in which signaling is the only role of college education. Allowing for heterogeneous effects of the reform (using Athey and Imbens' (2006) changes-in-changes estimator), I find a homogenous impact along the wage distribution; this indicates that wages declined proportionally for high- and low-earners.

I investigate the mechanisms that led to lower wages. Using data for economics graduates from Los Andes, I find that the distribution of employers changed with the reform, and that the likelihood of being employed by the highest-paying firms decreased. Moreover, I find that there is a relationship between the classes dropped and the placement of graduates across employers. Using data from the recruitment process for economists at the Central Bank from 2008 to 2014, I find that for graduates of Los Andes, the probability of being hired fell by 17 percentage points after the reform. This suggests that the reduction in courses introduced by the reform, decreased students' performance in recruitment processes, which in turn placed them in lower-quality firms and ultimately decreased their wages. Given that initial firm placement plays a significant role in determining long-term labor market success (Oreopoulos, von Wachter and Heisz, 2012), my results could also hint at possible longer-term effects. These results, however, are not estimates of the internal rate of returns to investment in additional schooling, but simply the effect on wages early in people's careers.

Finally, I examine possible threats to my identification strategy. First, it could be that the curriculum reform changed the pool of applicants and entrants in dimensions that are not captured by the Saber 11, but are relevant to the labor market. Specifically, given the decline in requirements for graduation, lower-ability individuals should be induced to enroll in these programs, which would lead to a decrease in the value of the signal and in wages. To address this

concern, I estimate an alternative specification, taking as the treatment group students at Los Andes who were already enrolled at the time of the reform but studied under the new curriculum. Results for this alternative treatment group are similar to the baseline specification. Second, my estimates might capture a negative trend in the return to a degree from Los Andes. To test whether this is the case, I perform two exercises: First, I replicate my baseline estimation using a placebo date for the reform, and second, I test my specification using a major at Los Andes that did not undergo a curriculum reform. I do not find evidence of wage changes in either case. My results are robust to several additional checks explained in the robustness section. Lastly, to interpret the reduction in wages as the causal effect of human capital, the choices underlying labor force participation should be unaffected by the reform. To check that this is the case, I estimate the effect of the reform on the probability of being employed in the formal sector. I find that for both economics and business the effects is very close to zero and statistically insignificant.

This paper contributes to the literature by estimating the effect of human capital accumulation on wages, separate from that of signaling, in a college setting. To the best of my knowledge, this is also the first paper to investigate the mechanisms that led to changes in wages; as a result, the study provides important information about the tools employers use to learn about workers' expected productivity.

A number of papers have investigated this issue in primary and secondary education settings and obtained mixed results. Eble and Hu (2016) exploit the introduction of one extra year in primary school in China in 1980 and find a 2% increase in wages. Since this accounts for a small fraction of the overall return to schooling, they conclude that there is an important role for signaling in primary education, however, no extra coursework was introduced in that additional year. Lang and Kropp (1986) and Bedard (2001) find secondary schooling decisions

that are consistent with a signaling model and would reject a pure human capital framework. Another strand of the literature attempts to directly measure whether there is a signaling value to academic degrees. Tyler, Murnane, and Willett (2000) estimate the signaling value of the GED to be between 12% and 20%, whereas Clark and Martorell (2014) find little evidence of high school diploma signaling effects.

Finally, my results suggest that human capital accumulation is an important explanation for the returns to college education. This finding relates to a growing literature that estimates the returns to different types of post-secondary education by separating the effect of the institution versus the field of study. This literature suggests that what matters the most is the type of degree as opposed to the institution from which it was obtained (Dale and Krueger 2002, Kirkeboen et al. 2016, and Hastings et al. 2013). This paper contributes by providing evidence that suggests that the source of heterogeneity in returns may be due to the different sets of skills and knowledge acquired in each degree, and not to differences in selectivity.

The rest of the paper is structured as follows. Section 2 describes a simplified version of a signaling and human capital model to derive testable implications in my context. Section 3 discusses the curriculum reform at Los Andes, and Section 4 describes the data, empirical strategy, and results. In Section 5, I explore the mechanisms that explain my results. Section 6 presents robustness checks, and Section 7 concludes.

1.2 Theoretical framework

In this section I lay out a simple model that allows me to derive a test of the signaling and human capital theories by exploiting a curriculum reduction at a top university, in a context of ability-based admissions and a binding number of slots.

Individuals have ability θ_i distributed with continuous support. There are J schools that offer different levels of human capital accumulation f_j , where j indicates school ranking. The cost to attend school j for individual i increases with the level of human capital and decreases with the level of ability, such that $c(f_j, \theta_i) > c(f_k, \theta_i)$ for every i when $j < k$, that is, when j offers higher human capital than k , and $c(f_j, \theta_i) < c(f_j, \theta_m)$ when $\theta_i > \theta_m$. In addition, $\frac{\partial^2 c(f, \theta)}{\partial f \partial \theta} < 0$, meaning that the cost of attending harder schools increases less for higher-ability individuals.

Suppose that productivity is a linear function of ability and human capital such that for a given set of beliefs regarding the assignment of students to colleges, expected productivity takes the form in equation (1).

$$w_j = \mu(E[\theta_i | f_j], f_j) = \alpha_1 + \alpha_2 \bar{\theta}_j + \alpha_3 f_j \quad (1)$$

These beliefs are written $E[\theta_i | f_j] \equiv \bar{\theta}_j$, where $\bar{\theta}_j$ is decreasing in j (i.e., firms believe that average ability is greater in higher-ranked colleges), and firms observe the level of instruction f_j . In a separating Perfect Bayesian Equilibrium, agents signal their type, and firms predict ability based on the observed level of human capital and offer wages accordingly. Students choose the school j that maximizes wages net of effort costs:

$$w_j - c(f_j, \theta_i) = \mu(E[\theta_i | f_j], f_j) - c(f_j, \theta_i) \quad (2)$$

Thus, a student chooses to attend the top school whenever:

$$w_1 - w_2 \geq c(f_1, \theta_i) - c(f_2, \theta_i) \quad (3)$$

The left-hand side of (3) is a positive constant, whereas the right-hand side is decreasing in θ_i . Then, there exists a unique θ^1 such that $\forall \theta \geq \theta^1$ (3) will hold (see Appendix 1 for the proof and a simple example). Subsequently, there is a threshold θ for each pair of schools that determines school choice over the school's ranking.

In this framework, the question of signaling vs. human capital comes down to learning about the values of $\alpha_2 \bar{\theta}$ and $\alpha_3 f$ in (1). To identify the contribution of human capital to wages, we need variation in f that holds θ constant. If school No.1 reduces the quantity of human capital produced, ($\Delta f_1 < 0$), such that it is still higher than f_2 , this model would predict that since the effort required to attend school No.1 went down, the level of ability that determines for whom it is profitable to attend the best school would also decrease, and thus $\bar{\theta}_1$ would decrease, and the fall in wages will confound the effects of the decline in the average ability of students and the decline in learning: $\Delta w_1 = \alpha_2 \Delta \bar{\theta}_1 + \alpha_3 \Delta f_1$. Note, however, that in an environment in which school No.1:

- i) Is constrained to admit a certain maximum number of students.
- ii) Uses a proxy of ability to determine admissions.
- iii) The maximum number of students is binding before the curriculum change.

Then:

By selecting students based on test scores the admissions criteria guarantee that the quality of the admitted class will not be affected by the reform, because the school was already choosing a subset (i.e., those with highest ability) of the group of applicants who find it profitable to attend school No. 1.

And thus:

$$\Delta w_1 = \alpha_3 \Delta f_1 \quad (4)$$

If α_3 is zero, the data support a pure signaling model in which wages are solely determined by the school's average student ability; see equation (1). If α_3 is statistically different from zero, this suggests a role for human capital in the determination of wages. In the next section, I will review the assumptions that lead to this result.

1.3 Institutional background and reform

I will first describe the salient characteristics of Colombian education and labor market institutions and details of the curriculum reform. On the education front, college admissions occur twice a year. Students apply directly to a major, and the gross enrollment rate in higher education is around 39%. Regarding the labor market, recent graduates are typically recruited year-round, and only a few multinational companies have a formal recruitment season. Recruitment at this level usually consists of tests of specific knowledge, standard human resources selection tests, and interviews. Twenty-five percent of college graduates work in the informal sector. Los Andes is a private university, and is ranked first in Colombia.

1.3.1 Reform

In 2006, Los Andes unilaterally decided to reduce the coursework required to earn a degree in most of its majors.⁴ The reasons given for the reform were to move towards international standards of shorter college degrees and to encourage graduate study. Each department was autonomous in implementing the reform. In this paper, I exploit the reforms

⁴ Los Andes was the only school to implement this practice at the time.

implemented by the economics and business departments. These consisted solely of a reduction in required credits; in other departments, the change led to the complete overhaul of curricula. In economics, the curriculum was trimmed by 12 courses (20% of the total number of credits), which resulted in a median number of five courses per term instead of six. Specifically, the reform: (i) took six mandatory courses and change them to optional courses (Monetary Policy, Public Finance, Trade, Marxist Economics, Colombian Economic Policy, and Social Programs Evaluation); (ii) reduced the number of optional courses by four; (iii) combined two probability and statistics courses into one; and (iv) combined accounting and economic measurement courses into one. The business department eliminated Computer Programming, Simulations, and Microeconomics I. In addition, the requirement of six upper-division electives was reduced to three. For both majors, instruction time was reduced from 4.5 years to 4 years.

The reform affected new students and students who, at the time it was implemented, were beginning their second year or earlier for economics, and in their third year or earlier for business. Other enrolled students were not affected by the change.

1.3.2 First stage: Empirical evidence of the reform for economics and business

To separately estimate the effects of human capital from that of signaling, I need an effective decline in the number of terms studied and credits earned, with no changes in the quantity or quality of the pool of students graduating from Los Andes. To investigate these points, in this section I present data on aggregate statistics from Los Andes' annual bulletins and micro data on credits earned by economics students.

Was the reform effective?

Figure 1 shows the average duration of undergraduate programs for both economics and business majors. There is a step down in these trends of about one semester at the time of the reform, which suggests that the reform was effective in decreasing the average length of the program. For economics, the average duration went from 5 to 4.5 years, and for business duration declined from 5.5 to 5 years. **Figure 2** shows the number of credits students graduated with in economics. We can observe a sharp drop at the time of the reform of around 16%. **Table 1** shows regression estimates from fitting different linear trends around the reform.

Did the reform affect the size and composition of the entering and graduating classes?

To evaluate this question, I check the evolution of the size of the entering classes, their average Saber 11 scores, and average graduation rates. *Panel a* of **Figure 3** shows the evolution of the entering class in economics and business. I fit different trends before and after the reform. The graph shows that the number of entering students was not affected by the reform.⁵ *Panel b* of **Figure 3** shows the average Saber 11 scores of the entering class. Fitted regressions around the reform do not suggest a change in the quality of the entering class. I also perform a DID estimation, similar to the one I perform for my baseline analysis, to determine whether the reform reduced the average Saber 11 score. Specifically, I compare Saber 11 scores for incoming cohorts

⁵ Even though I do not find a discontinuity in test scores, there is a change in trends around the time of the reform. This could be problematic for my identification strategy if the control group does not experience the same change in trend as Los Andes. To check for this possibility, Figure A1.2 shows Saber 11 scores for entering cohorts at Rosario University and reveals a similar pattern.

before and after the reform for Los Andes and other Top 10 schools. **Table A2.1** shows that there is no evidence of a reduction in scores after the reform at Los Andes.

On the other hand, if the change in curricula alters the quantity of students *graduating* from Los Andes, the value of the signal would change. This is plausible, since the requirements to graduate decreased with the reform. *Panel c* of **Figure 3** shows the evolution of graduation rates and suggests that the reform had no effect on the dropout rate. I also perform a DID linear probability model regression to test whether the reform changed the probability of graduating with an economics or business degree, and do not find evidence that it did (**Table A2.1**). **Figure A3.1** also shows that the reform did not change the share of students who graduated with a minor.

In the model, students use school rankings to choose which college to attend; if the reform decreased Los Andes' ranking, post-reform cohorts would have, on average, lower ability. Even though the above doesn't provide evidence of this, I examine this point directly by looking at rankings and college exit exam scores. International rankings that include Latin American universities are only available since 2013, but from 2013 to 2016, Los Andes has been ranked as the best school in Colombia.⁶ The Colombian Ministry of Education released its first rankings in 2015, in which it also ranked Los Andes first.⁷

To summarize, the reduced curriculum translated into an effective cut of one semester from the average degree duration for economics and business and a reduction in the number of credits per term; this constitutes an exogenous reduction in human capital. On the other hand, the number of new students, Saber 11 scores, and dropout rates suggest that the quantity and quality

⁶ <https://www.timeshighereducation.com/world-university-rankings/2015/world-ranking#!/page/0/length/25>
<http://www.topuniversities.com/university-rankings/latin-american-university-rankings/2014#sorting=rank+region=+country=+faculty=+stars=false+search=>
Accessed February 10, 2016.

⁷ <http://www.mineducacion.gov.co/cvn/1665/w3-article-351855.html> Accessed February 10, 2016.

of students was unaffected, and therefore the selectivity of the degrees remained unchanged after the reform. This is an ideal environment to test the role of signaling and human capital in college education.

1.4 Effects of the Reform

In this section, I estimate the effect of the reduction of the curricula on wages to test the prevalence of a pure signaling model versus a model in which human capital matters. I start by describing my data, continue with the identification strategy, and end with the results.

1.4.1 Data

I use administrative data from the Ministry of Education. My main database is the OLE (*Observatorio Laboral de Educación*), which is constructed to follow yearly earnings in the formal sector for college graduates in Colombia.⁸ This information is recorded from Social Security payments from 2008 to 2012. The OLE also contains education variables, such as university and program attended, graduation year, and personal characteristics.

SPADIES (*Sistema para la Prevención de la Deserción en la Educación Superior*) is a database that tracks college dropout rates. Like the OLE, it contains data on university attended but also has information on the first semester of college, which I needed to identify each student's curriculum. This database also contains household socioeconomic variables. The third database contains individual data on Saber 11 scores.

⁸ 75% of workers with a college education are employed in the formal sector (Fedesarrollo, 2013).

The three databases contain generated ID numbers to trace individuals⁹. My baseline database contains all students who started college between 2002 and 2007 and graduated after 2004 from Los Andes and other top-10 schools, and who enrolled in economics or business. **Table 2** shows summary statistics of some relevant variables in the data. The average individual in my sample is 26 years old and has been working for almost three years¹⁰. On average, Los Andes graduates earn 45% more than graduates of the next 10 schools in the national rankings (“Top 10” hereafter) and have higher Saber 11 scores; their parents also have higher incomes.

1.4.2 Preliminary evidence and empirical strategy

Figure 4 shows a scatter plot of wages for graduates from Los Andes and Top 10 schools for economics and business by cohort. Before the reform, the evolution in wages seems fairly parallel, and the slopes for wages are statistically the same. There was a constant premium for attending Los Andes of 36% for economics and 50% for business. With the curriculum change, the premium immediately declined for economics and gradually for business, for a final average reduction of 22 percentage points and 12 percentage points, respectively. **Figure A3.3** displays the wage densities for Los Andes and the Top 10 schools, both before and after the reform. The graphs show that for the control group, pre- and post-reform wage densities overlap each other, but for Los Andes, post-reform densities shift to the left. Both **Figures 4** and **A2.3** show that the reform had a starkly negative effect on the wage distribution of Los Andes graduates. To estimate the magnitude of human capital’s role in wages, I estimate the following DID regression:

⁹ Anonymized identifiers are generated using national identification numbers, name, and date of birth.

¹⁰ The fact that my data consist of wages from individuals at the beginning of their professional careers poses a challenge to my specification, since wage profiles are very steep in terms of experience.

$$\ln wage_{it} = \beta_0 + \beta_1 Andes_i * Post_t + \beta_2 Andes_i + \beta_3 Post_t + \beta_4 experience_{i,t} + \varepsilon_{it}, \quad (1)$$

where $wage_{it}$ is the average monthly earnings of student i in year t , (in 2010 pesos). *Andes* is a dummy equal to 1 if student i went to college at Los Andes, and 0 if he went to another Top 10 Colombian university (my baseline control group). *Post* is a dummy equal to 1 if a student started school after the date of the reform implementation and 0 otherwise, and experience is measured in years since graduation. The coefficient β_1 captures the effect of graduating from Los Andes after 2006 on wages. I also control for gender, year, and cohort effects in other specifications.

1.4.3 Results

Tables 3a and 3b show my baseline results: *Panel a* presents estimates for economics and *panel b* for business. The baseline estimation for equation (1), reported in column 1, indicates a decline in wages by 16% for economics and 13% for business. Column 2 adds controls for experience squared and gender, and columns 3 through 6 add year and cohort controls to these specifications. Throughout all such specifications, there is a negative and strong decline in wages as a result of the reform. These results reject a pure signaling model, in which wages should not change; given the magnitude of the decline, they demonstrate an important role for human capital in the determination of wages.

While estimation is straightforward in this setting, statistical inference is not. Moulton (1990) shows that the failure to account for the presence of common group errors leads to insufficiently conservative inference. In response to this concern, a clustering procedure emerged

whereby inference relies on the asymptotic approximations associated with the assumption that the number of individuals within a group and/or the number of groups grows large. However, this assumption does not apply in my setting. To address this concern, I follow Abadie, Diamond and Hainmueller (2010).¹¹ Their inferential exercise examines whether the estimated effect of the actual intervention is large relative to the distribution of the effects estimated for schools that are not affected by the reform. To implement this procedure, I estimate equation (1) an additional 10 times, replacing Los Andes with an indicator for one of the other 10 schools. For all cases, the estimate for Los Andes was the largest (**Table 4**). I also evaluate equation (1), changing the date of the reform, in addition to the treated school. **Figures 5-7** show the distribution of treatment effects: In all specifications the effect I estimate is at the 5th percentile mark, or to the left.

With the data available, I can only estimate the effect of the reform on earnings early in students' careers. However, it is at this stage that the debate over signaling and human capital is particularly relevant, given that with time, employers learn about students' productivity on the job (Farber and Gibbons (1996), Altonji and Pierret (2001) and Lange (2007)). Specifically, consistent with my findings, Lange (2007) finds that employers learn quickly; initial expectation errors decline by 50% within 3 years. Lange also estimates that signaling contributes less than 25% to gains from schooling.

It is possible that the reform changed the pool of applicants and entrants in dimensions not captured by the Saber 11 that are relevant to the labor market. Specifically, given the decline in requirements to graduate, lower-ability individuals should be motivated to enroll in these programs—thereby decreasing the value of the signal and, in turn, wages. To address this, I

¹¹ See also Imbens and Wooldridge (2009) and Cunningham and Shah (2017) for recent applications of this method.

estimate an alternative specification in which the treatment group consists solely of Los Andes students who were already enrolled at the time of the reform, but studied under the new curricula. **Table 5** and **Figure 6** show results for this alternative treatment group. According to the data, there is a strong and negative effect on wages of around 16% for economics and 12% for business.

Given that the number of years of wage observations by group is unbalanced (pre-reform vs. post-reform and treated vs. untreated), in **Table 6** and **Figure 7** I include observations with at most three years of experience, to ensure that the treatment coefficient is not capturing differences in the slope of the experience profile. Results in **Table 6** again suggest strong wage declines of the same magnitudes as those found previously.

To make use of all data available, and recognizing the potential for heterogeneous effects, I now turn to a changes-in-changes (CIC) estimation following Athey and Imbens (2006) and Melly and Santangelo (2015) that extends the model to include covariates. I estimate CIC for the 10th through 90th percentiles after controlling for experience, gender, and cohort effects. As can be seen in **Figure 8**, there is little evidence of heterogeneity in the reform's effect on wages by percentiles and fields, suggesting that the assumptions of the traditional DID estimator hold.

The decline in wages I find is large, and suggests sizable estimates of the return to attending college at Los Andes. To better understand the magnitude of my estimates, I perform a back of the envelope calculation that attempts to quantify the reduction in the wage premium of attending college at Los Andes. Using a cross-section estimate of the return to college and an estimate of the return to Los Andes from Saavedra (2008), I find that the return to attending college at Los Andes relative to not attending college is 110% (details of this calculation are explained in Appendix 2). This implies that the reform reduced the premium by 14.5% and 11.8%

as a result of a reduction in credits of 20% and 14% in economics and business, respectively. However, in the absence of a causal estimate of the return to college for this setup, I cannot decompose this return.

1.5 Mechanism

When and how do employers find out about these graduates' lower human capital? Specifically, were they able to detect it in the recruitment process, during tests or interviews? Or did they notice it on the job? Unfortunately, I do not have the information necessary to fully answer these questions, but I do have data collected by the economics department at Los Andes, on the employers of all economics graduates by cohort, which I use to investigate whether employers changed with the reform. **Table A3.3** lists the main employers before and after the reform, and shows that there are important differences. There seems to be a connection between the change in curriculum and the change in employers: The Central Bank, the Ministry of Finance, and the National Planning Department are less likely to employ economists who graduated under the new curriculum, under which the courses Monetary Policy, Public Finance, and Colombian Economic Policy were no longer mandatory. Indeed, Figure A3.6 shows that there was a decline in the number of students enrolled in these classes after the reform. From this comparison, I also find that the likelihood of being employed by the highest paying firms decreased with the reform. Using a ranking of the 100 highest paying firms for recent graduates in economics, I find that the share of students in these firms fell from 24% to 14% after the reform.

I interviewed employers to learn about their experiences with hiring economics graduates, and as anecdotal evidence I learned that:¹² (i) most knew about the reform from talking to recent graduates; (ii) they believe they can detect changes in human capital through tests they administered in the recruitment process; (iii) they argue that for some jobs, the content made optional in the new curriculum is critical; (iv) they believe that taking fewer elective courses affects graduates' labor prospects beyond the recruitment process, because the professors in those courses are helpful with job offers and job referrals; and (v) wages for new graduates are fixed. All of the above provides suggestive evidence that under the new curricula, the pool of jobs a graduate can obtain is smaller, either because they cannot succeed in the recruitment process—which includes tests on content they did not cover in school—or because they have less contact with professors who have connections in the job market. It is clear that the first reason is entirely due to a decrease in human capital, but this is not the case with the second.

To evaluate whether the reform had an impact on students' ability to obtain jobs, I perform a DID exercise with data from the recruitment process for recently graduated economists at the Central Bank of Colombia. This consists of a written exam, which tests specific knowledge necessary for the position, as well as human resources tests and interviews with both human resources staff and department heads. Most such openings are publicly announced through employment websites and social networks, and are open to any and all applicants. I have data on university and enrollment terms for all candidates for economist positions from 2008 to 2014, along with the final hiring decision. For candidates who studied under the old curriculum, the probability of being hired was 27%; this fell to 6% with the reform. **Table 7** shows the results of

¹² I conducted interviews with 11 out of the 14 employers listed on the left panel of Table A2.3.

the DID exercise. According to data, after the reform there is a reduction of 16.7 percentage points in the probability of being hired by the Central Bank for students from Los Andes versus students from Top 10 schools. This suggests that one of the possible mechanisms that led to the decline in wages is a decline in the performance of students during the recruitment process. As a result, the pool of offers a student could choose from was smaller, and students started in lower-paying jobs.

The previous mechanism points to an environment in which employer learning happens rapidly due to the availability of tests on specific content used in the recruitment process. Thus, one would expect that employers notice the reduction in instruction at Los Andes soon after the first students enter the job market, which is what happened in economics (**Figure 4**). For business, however, if the recruitment process relies less on testing specific knowledge, we would not expect to see this pattern. Interviews with recruitment agencies suggest that this is the case, since a large share of the openings for recent graduates in business are also available to graduates from economics and engineering, and thus tests on content are less appropriate. As a consequence, the qualitative evidence suggests that it might take longer for employers of business students to notice the differences in human capital of the new cohorts.

1.6 Robustness Checks

In this section I perform several robustness checks to address possible confounding factors in my estimation. I then discuss some important caveats and limitations.

It is possible that my estimates capture a negative trend in the return to a degree from Los Andes. To determine whether this is the case, I replicate my baseline estimation using a placebo

date for the reform. Specifically, I include only cohorts that studied under the old curricula, and set a fake reform date in the middle of the period covered. If my results were driven by a decline in the return to Los Andes, any *post*Andes* interaction would be negative and statistically significant. However, as shown in **Table 8**, all of the estimated effects are statistically equal to zero and smaller than 0.7% in economics, and positive for business.

An alternative placebo check to address this concern is to test what happens to law graduates (a major whose curriculum was not reformed) during the dates of the reform in economics and business. Results in **Table 9** show that there is no effect on wages for Los Andes law graduates on the date of the reform in economics or business. All of the above suggest that the strong decline in wages I find is not the result of other trends or changes at Los Andes.

Table 10 presents a series of additional robustness checks. The first two columns show results for economics and the last two for business; columns 1 and 3 estimate equation 1 with cohort controls, and columns 2 and 4 add experience squared and gender. A possible explanation for these results is that there is an age penalty in the labor market. We can imagine that if two graduates have the same credentials, employers might lean toward the older one, thinking that life experience is valuable for the job. In this case, having cohorts that graduate half a year younger would result in lower wages, regardless of human capital or signaling considerations. To check this possibility, I include age as an independent variable in my baseline estimation. The results in *panel a* of **Table 10** suggest that there is a strong effect of the reform outside of age considerations. For economics, the effect is the same (-16%), and for business it is smaller (-9%).

One might also be worried about the fact that the reform generated two cohorts that graduated at the same time, which might have distorted wages by creating more competition. In

panel b of **Table 10**, I exclude these two cohorts and perform my baseline estimation; results show that the effects hold, even with the exclusion.

An additional concern about the previous estimates is the validity of the control group. Even though the pre-trends in wages were similar, the control group might not be a good counterfactual—if, for example, the two groups face different labor markets, and these evolved in different ways after the reform. To address this, I limit my control group to students graduating from the next three highest ranked schools, because it is likely that students from these institutions will face the same labor market as students from Los Andes. *Panel c* of **Table 10** presents the results of the reform’s effect on wages under this alternative control group; we can see that there is a negative effect of the reform on wages of similar magnitude to the one found before.

An alternative way to address the concern about the validity of the control group is to include only students who had the academic credentials required to attend Los Andes in the control group. Specifically, I include students who attended Top 10 schools and had Saber 11 scores greater than the minimum per cohort observed at Los Andes for economics and business. *Panel d* of **Table 10** shows the results of this alternative exercise: Wages fall by a magnitude larger than in the baseline estimation (18% for economics and 15% for business).

Panel e of **Table 10** repeats the baseline estimation, excluding cohort 2007-1; as shown in **Figure 4**, this cohort had particularly low wages for students from Los Andes. Again, the results are very similar, suggesting strong declines in wages. Finally, *Panel f* includes Saber 11 scores as a covariate. We can see that when controlling for test scores, the results hold and even increase slightly.

Since there are multiple possible choices for control groups, I follow Abadie and Gardeazabal (2003) and perform a synthetic control exercise in which I look for the best combination of major and school to match the pre-trend data of my treated groups. The comparison unit in the synthetic control method is selected as the weighted average of all potential comparison units that best resembles the characteristics of the case of interest. **Table 11** shows the results of my baseline specification with respect to the optimally chosen control group. This group features engineering, business, and law graduates of Top 10 schools. Using this method, results are similar to the ones found previously: The reform's effect for economics graduates ranges from -7% to -13%, and for business graduates there is a larger dispersion, with the effect ranging from -5% to -20%.

Finally, in the previous analysis I assumed that the reform did not have an effect on labor force participation. To check that this is the case, I estimate the effect of the reform on the probability of being employed in the formal sector. Table A3.3 shows the results of a DID regression on the probability of being employed. For both economics and business the effects is very close to zero and statistically insignificant.¹³ In addition, since one of the motives for the reform was to increase graduate school enrollment, it is important to check for changes along this dimension. It is possible, for instance, that before the reform only students in the right tail of the ability distribution attended graduate school, but after the reform more students enrolled, and therefore the estimated difference in wages results from comparing wages from different segments of the ability distribution. To determine whether this is the case, I use LinkedIn and personal and firm websites to obtain information on graduate school enrollment for the last three

¹³ Saavedra (2008) finds a positive effect of attending Los Andes on employment (p. 22, table 8). The first difference of the regression on Table A3.3 supports this finding.

cohorts that studied under the old curriculum and the first three that studied under the new one. **Figure A3.4** shows that the percentage of graduates found on LinkedIn—around 60%—is similar to the rates before and after the reform. **Figure A3.5** also shows the share of graduates by cohort who enrolled in graduate school in the first four years after obtaining an undergraduate degree, and the shares do not seem to increase with the reform. All of the above suggests that selection does not appear to be driving the decline in wages.

1.7 Conclusions

In this paper I identify the effect of human capital on wages by exploiting a curriculum change at Universidad de los Andes in Colombia. In 2006, the amount of coursework required to earn a college degree in economics and business decreased by 20% and 14%, respectively. This was accomplished by dropping 12 courses in economics and 6 in business, and a reduction in instruction time of one semester. The reform did not alter the quality of the entering or graduating classes or the school ranking. Because wages should fall under the human capital model—but remain constant under pure signaling—this constitutes an ideal natural experiment for learning about signaling vs. human capital.

Using administrative data on wages and college attendance from 2008 to 2012, I find that wages fell by 16% in economics and 13% in business. Given the size and statistical significance of the decline in wages, my estimates suggest that human capital plays an important role in the determination of wages. The results also reject a model in which signaling is the only function of college education. Note that this result does not rule out completely a role for signaling. For

example, also using data on Colombia, Macleod et al. (2017) find evidence of a signaling role in college reputation.

I use data and interviews from employers of economics graduates to study the mechanisms that led to the decline in wages. I find that the distribution of employers changed with the reform, and that the likelihood of being employed by the highest paying firms decreased. Employers argue that some of the content that was made optional in the new curricula was essential to the positions they offered; if that was the case, employers would have noticed that students had less human capital through knowledge tests in the recruitment process. This suggests that under the new curricula, the pool of jobs a graduate can obtain is smaller because they perform worse during the recruitment process, which subsequently decreases their wages. Using recruitment data from the Central Bank, I find support for this hypothesis and estimate that the reform reduced the probability of being successful by 17 percentage points.

1.8 References

- Abadie, A. and Gardeazabal, J., 2003. The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, pp.113-132.
- Abadie, A., Diamond, A. and Hainmueller, J., 2010. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association*, 105(490), pp.493-505.
- Altonji, J.G. and Pierret, C.R., 2001. Employer Learning and Statistical Discrimination. *The Quarterly journal of economics*, 116(1), pp.313-350.
- Athey, Susan, and Guido W. Imbens. 2006. "Identification and Inference in Nonlinear Difference-in Differences Models." *Econometrica* 74.2: 431-497.
- Becker, Gary S. 1962. "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy* 70, no. 5, pt. 2 (October): 9-49
- Bedard, Kelly. 2001. "Human capital versus signaling models: university access and high school dropouts." *Journal of Political Economy* 109.4: 749-775.
- Buchmueller, T.C., DiNardo, J. and Valletta, R.G., 2011. The effect of an employer health insurance mandate on health insurance coverage and the demand for labor: Evidence from hawaii. *American Economic Journal: Economic Policy*, 3(4), pp.25-51.
- Cameron, A.C. and Miller, D.L., 2015. A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2), pp.317-372.

- Card, David. 1999. "The causal effect of education on earnings." Handbook of Labor Economics: 1801-1863.
- Clark, Damon, and Paco Martorell. 2014. "The signaling value of a high school diploma." Journal of Political Economy 122.2: 282-318.
- Cunningham, S. and Shah, M., 2017. Decriminalizing indoor prostitution: Implications for sexual violence and public health. *Forthcoming* Review of Economics Studies.
- Dale, S.B. and Krueger, A.B., 2002. Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. The Quarterly Journal of Economics, 117(4), pp.1491-1527.
- Eble, Alex, and Feng Hu. 2016 "Demand for Schooling, returns to schooling and the role of credentials." Mimeo, Brown University.
- Farber, H.S. and Gibbons, R., 1996. Learning and Wage Dynamics. The Quarterly Journal of Economics, pp.1007-1047.
- Feng, Andy, and Georg Graetz. 2015. "A question of degree: the effects of degree class on labor market outcomes." No. 8826. IZA Discussion Papers.
- Hastings, J., Neilson, C. & Zimmerman, S. 2013. 'Are some degrees worth more than others? Evidence from college admissions cutoffs in Chile', NBER Working Paper 19241.
- Jepsen, Christopher, Peter R. Mueser, and Kenneth R. Troske. 2012. "Labor-market returns to the GED using regression discontinuity analysis." Working paper.
- Kirkebøen, L., Leuven, E. and Mogstad, M., 2016. Field of study, earnings, and self-selection. *Forthcoming*. Quarterly Journal of Economics.

- Lang, Kevin, and Kropp, David. 1986. "Human Capital versus Sorting: The Effects of Compulsory Attendance Laws." *Quarterly Journal of Economics*. 101 (August): 609-624.
- Lange, F., and R. Topel. 2006. "The Social Value of Education and Human Capital." In *Handbook of the Economics of Education*, vol. 1, edited by E. Hanushek and F. Welch. Amsterdam: Elsevier.
- Lange, Fabian, 2007. The speed of employer learning. *Journal of Labor Economics*, 25(1), pp.1-35.
- MacKinnon, J. G., & Webb, M. D. (2016). Wild bootstrap inference for wildly different cluster sizes. *Journal of Applied Econometrics*.
- MacLeod, W. Bentley, Evan Riehl, Juan E. Saavedra, Miguel Urquiola. 2017. "The Big Sort: College Reputation and Labor Market Outcomes". *Forthcoming: American Economics Journal: Applied Economics*.
- Mincer, J.A., 1974. *Schooling, Experience, and Earnings*. NBER Books.
- Melly, Blaise, and Giulia Santangelo. 2013. "The changes-in-changes model with covariates." Vortrag auf der Statistischen Woche.
- Montenegro, C.E. and Patrinos, H.A., 2014. Comparable estimates of returns to schooling around the world. *World Bank policy research working paper*, (7020).
- Moulton, Brent R. 1990. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Unit." *Review of Economics and Statistics*, 72(2): 334-38.

Oreopoulos, P., & Petronijevic, U. 2013. Making college worth it: A review of research on the returns to higher education (No. w19053). National Bureau of Economic Research.

Oreopoulos, P., Von Wachter, T. and Heisz, A., 2012. The short-and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics*, 4(1), pp.1-29.

Spence, A. Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics*. 87 (August): 355-74.

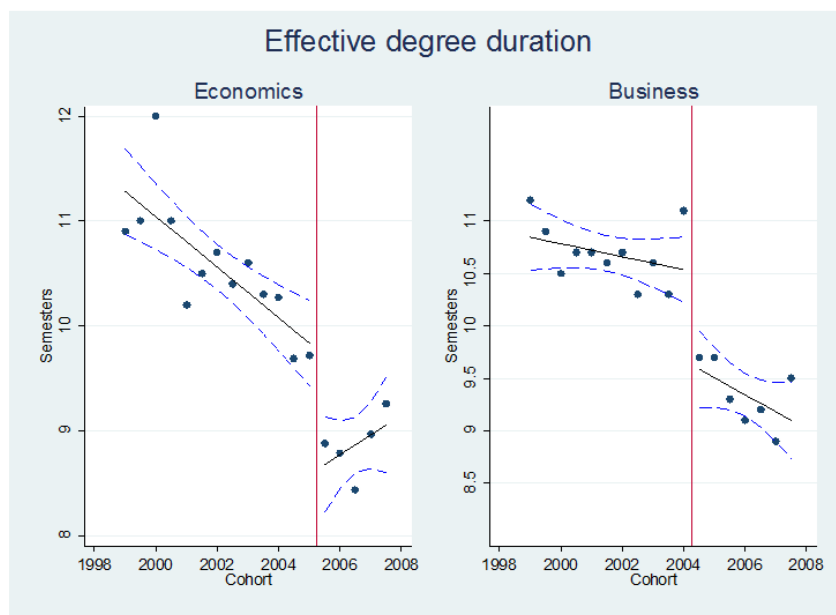
Tyler, J., R. Murnane, and J. Willett. 2000. "Estimating the Labor Market Signaling Value of the GED". *Quarterly Journal of Economics*, 115, 431-468.

Saavedra, J.E., 2008. The returns to college quality: A regression discontinuity analysis. Mimeo.

Weiss, A. 1995. "Human Capital vs. Signalling Explanations of Wages." *Journal of Economic Perspectives*, 9, 133-154.

1.9 List of Figures

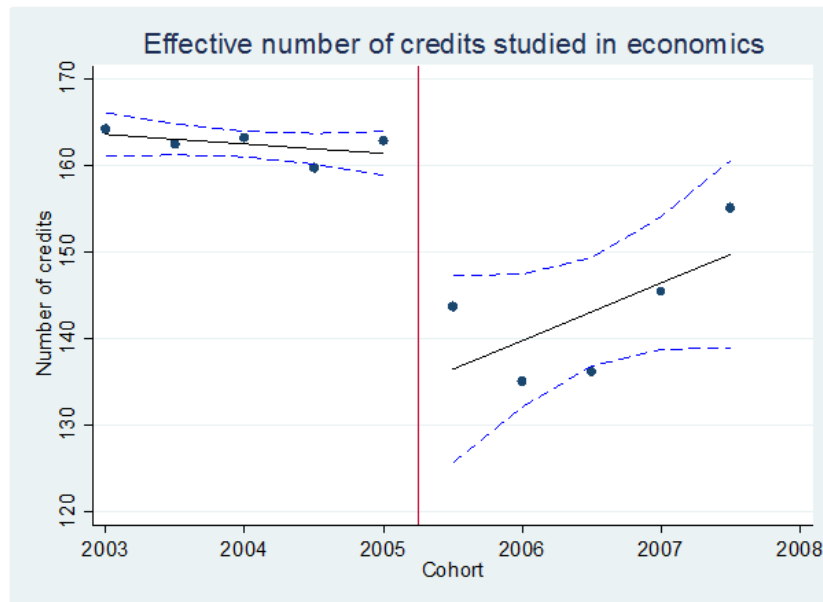
Figure 1.1: Effect of the reform in degree duration



Source: Annual statistical bulletin – Universidad de los Andes. Scatter plots are mean degree duration per cohort.

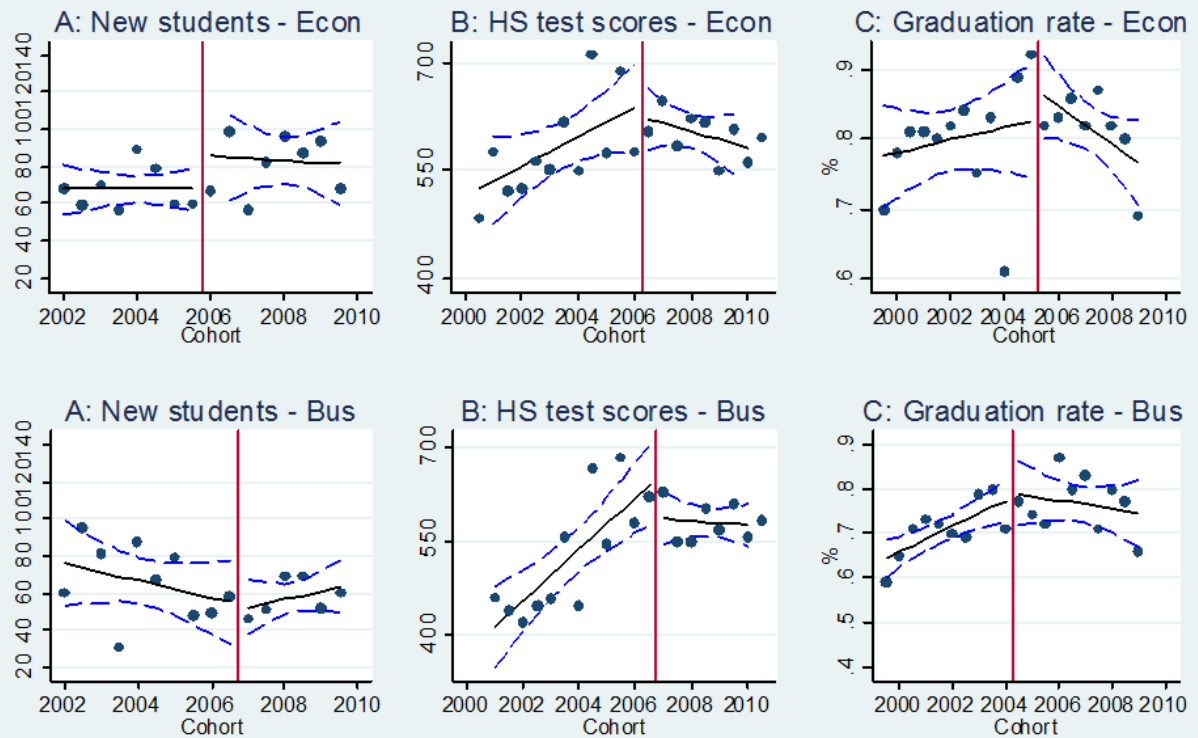
A cohort is defined by the semester the student started school. This graph includes all students who started the program. Solid lines are the fitted values of a linear regression on time, and dashed lines represent 95% CI of the estimation. The vertical line represents the time of the reform.

Figure 2: Effect of the reform in credits studied



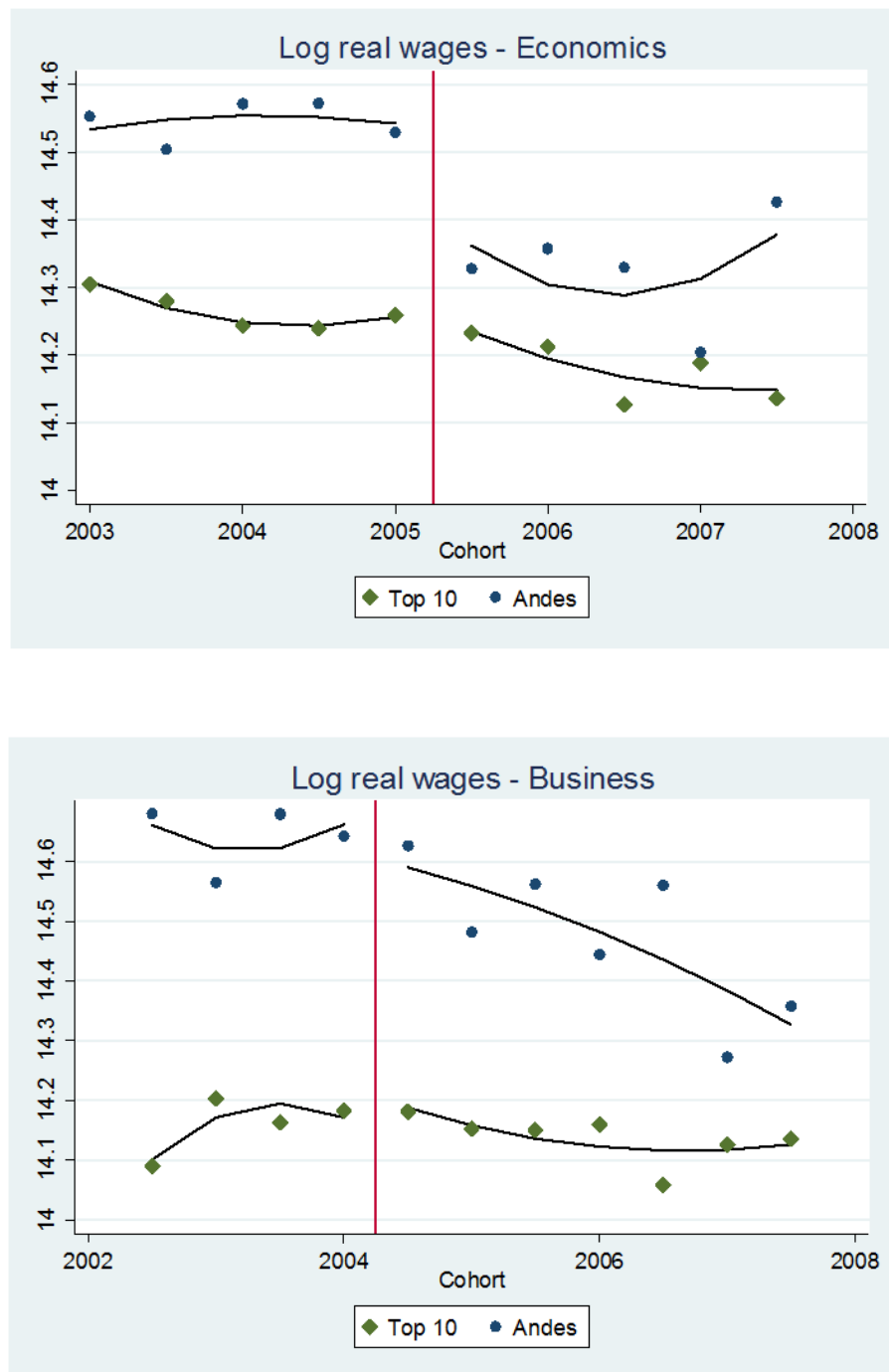
Source: Department of Economics – Universidad de los Andes. Scatter plots are credits studied by cohort. Solid lines are the fitted values of a linear regression on time and dashed lines are the 95% CI of the estimation. The vertical line represents the time of the reform.

Figure 3: Effects of the reform on class selection



Source: Annual statistical bulletins - Universidad de los Andes. The solid lines are the fitted values and dashed lines the 95% CI.

Figure 4: Pre trends and the effect of the reform on wages



Source: Ministry of Education. Scatter plots are mean wages per cohort and school group. Lines are the fitted values of a regression quadratic on time. The vertical line represents the time of the reform.

Figure 5: Treatment effect distribution (Table 3a and 3b)

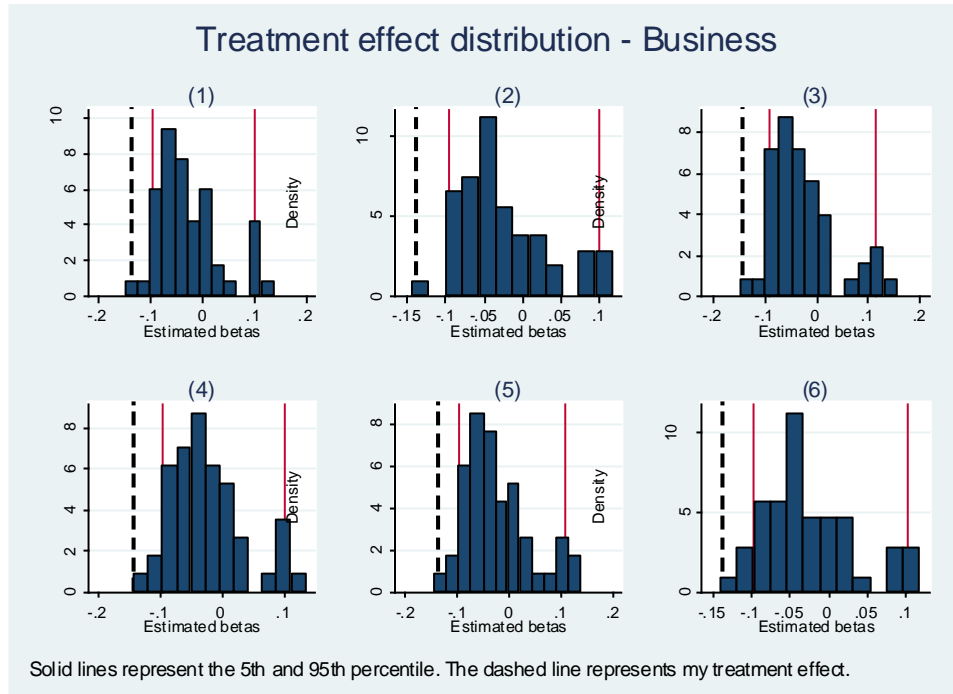
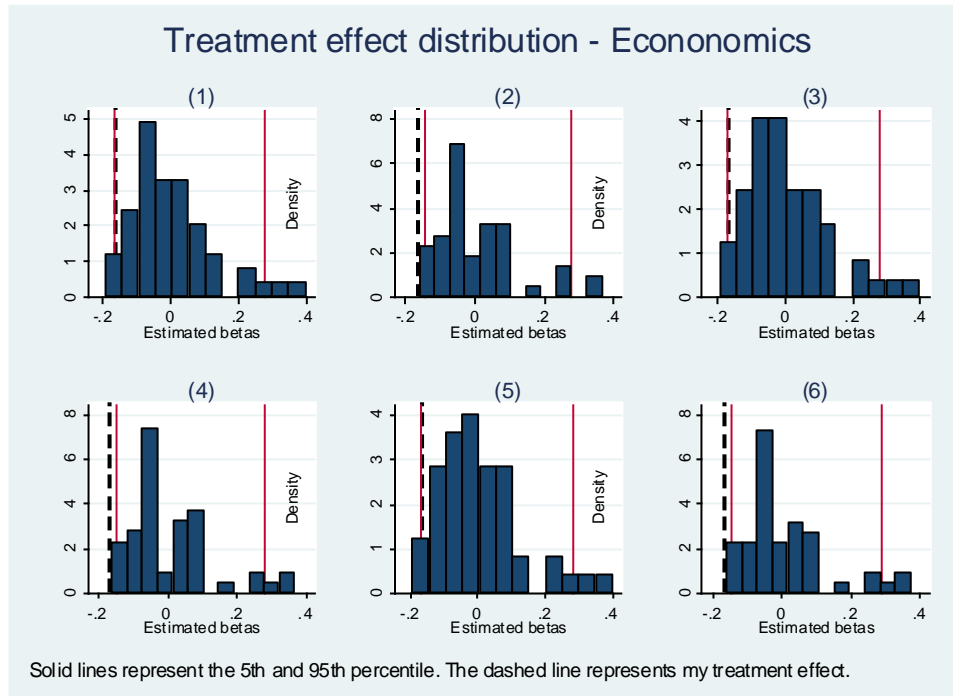
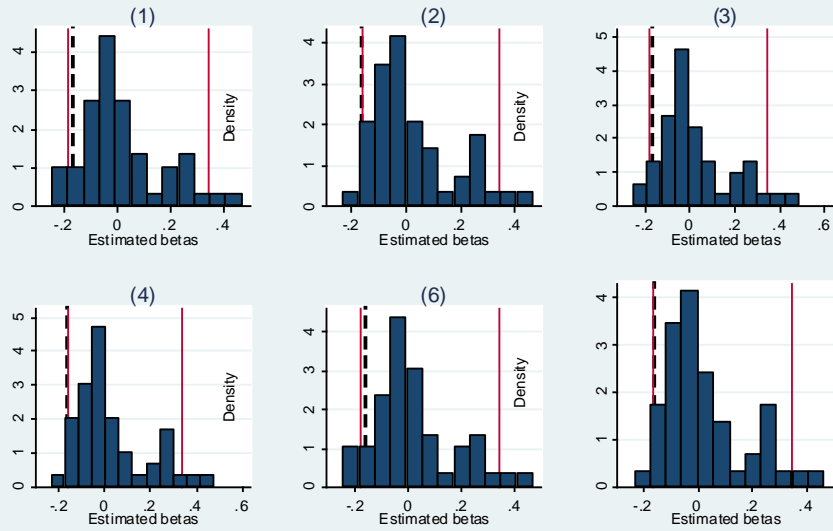


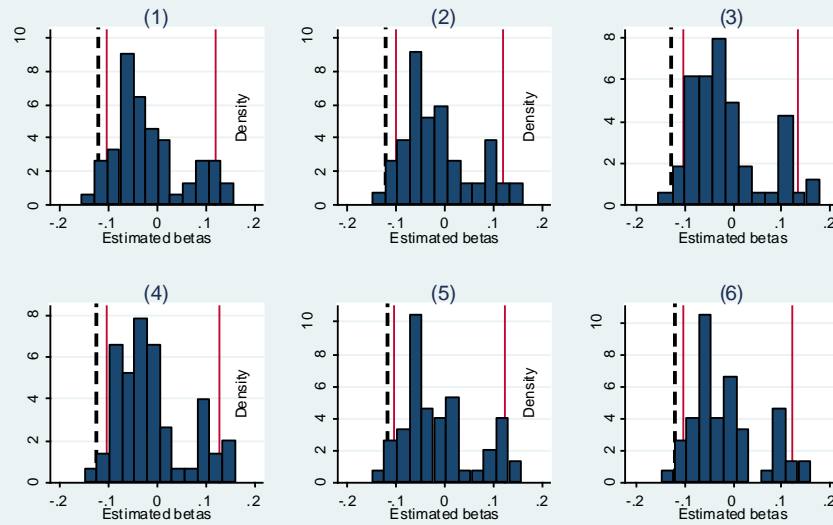
Figure 6: Treatment effect distribution (Table 5)

Treatment effect distribution (Alt. treatment group)- Economics



Solid lines represent the 5th and 95th percentile. The dashed line represents my treatment effect.

Treatment effect distribution (Alt. treatment group)- Business



Solid lines represent the 5th and 95th percentile. The dashed line represents my treatment effect.

Figure 7: Treatment effect distribution (Table 6)

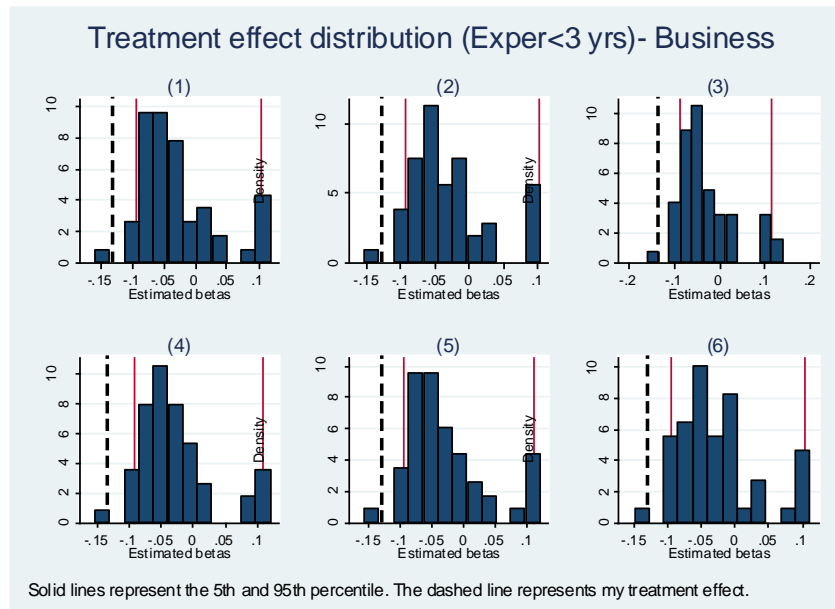
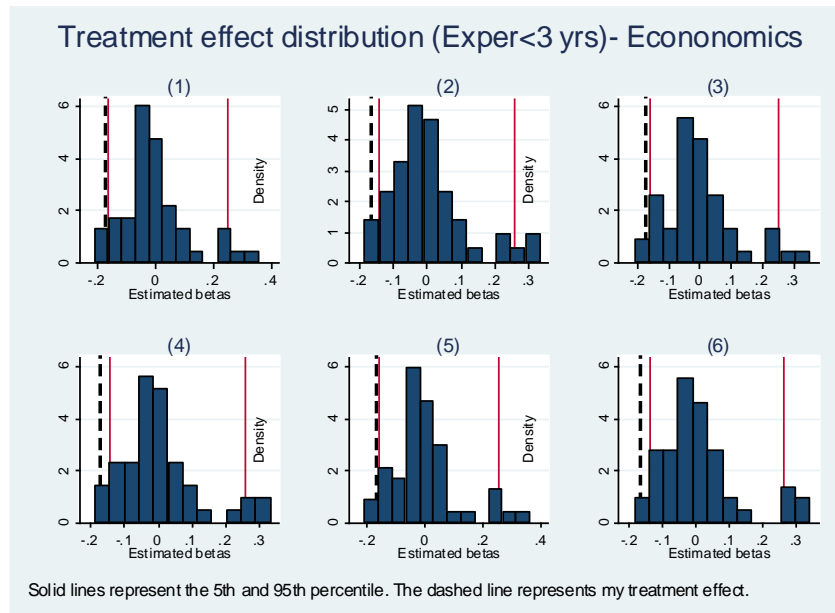
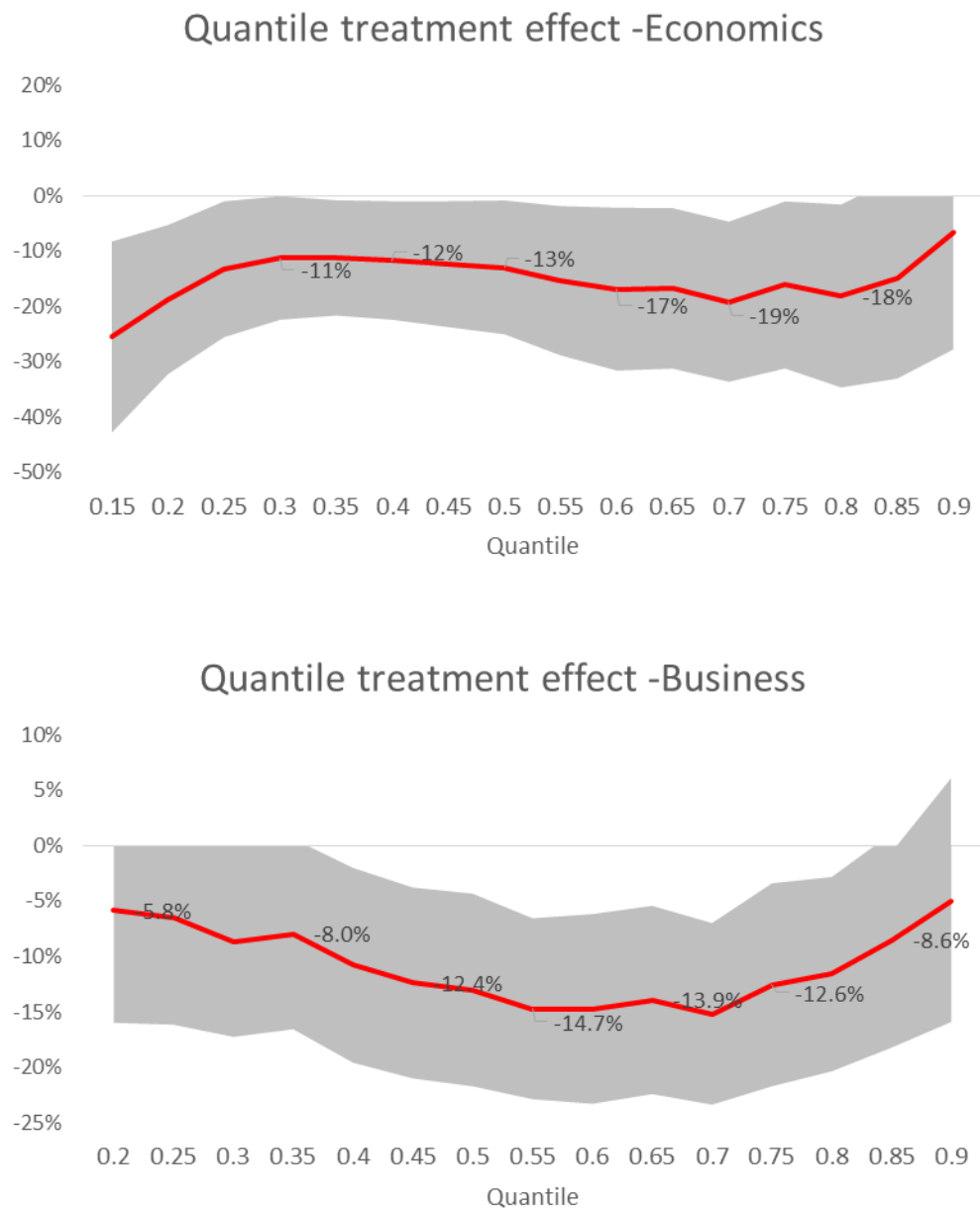


Figure 8: Changes in changes estimates



Source: Ministry of Education. CIC estimates of an estimation that controls for experience, gender, and cohort variables. Confidence intervals at the 95th percent level.

Test—Economics: Constant effect: $QTE(x)=QTE(0.5)$ for all x ; KS-statistic: 0.61; CMS-statistic: 0.47. Test—

Business: Constant effect: $QTE(x)=QTE(0.5)$ for all x ; KS-statistic: 0.374; CMS-statistic: 0.276.

1.10 List of Tables

Table 1: First stage –The effect of the reform on instruction and class quality

Dep variable:	Degree duration		No. of credits	Class size		HS test scores		Graduation rates	
	Econ	Buss	Econ	Econ	Buss	Econ	Buss	Econ	Buss
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post	-1.038**	-0.916***	-24.37**	8.56	-22.15	-1.396	36.42	0.0192	0.00097
	-0.367	-0.262	-6.751	-14.74	-32.2	-40.97	-71.51	-0.0635	-0.0496
Trend pre	0.0943	-0.0821	3.317*	1.024	2.086	-6.272	-1.655	-0.0135	-0.00503
	-0.119	-0.0528	-1.637	-2.248	-4.124	-5.546	-7.913	-0.011	-0.00606
Trend post	-0.121***	-0.0309	-0.545	0.0952	-2.309	12.90***	19.93***	0.00692	0.0145**
	-0.028	-0.0266	-1.637	-2.248	-1.899	-3.755	-3.801	-0.00596	-0.00606
Constant	9.716***	10.51***	160.8***	68.08***	64.35***	637.6***	557.0***	0.842***	0.789***
	-0.222	-0.181	-5.43	-9.403	-5.537	-21.28	-16.13	-0.0438	-0.0376
Obs	18	18	10	16	16	21	21	20	20
R squared	0.868	0.881	0.867	0.233	0.171	0.45	0.672	0.163	0.427

Source: Annual bulletin -Universidad de los Andes & Department of economics.

Table 2: Summary statistics

	Real wage	Experience	Age	Female	HS test	Family income*	Obs
Andes Economics	3,017,001	2.6	25.8	0.46	58.1	5.93	1,736
	1,776,674	1.9	2.2	0.50	5.5	1.44	
Top 10	2,119,275	2.98	26.26	0.59	51.28	3.75	3,580
	1,457,070	1.98	2.83	0.49	6.01	1.76	
Andes Business	3,192,033	2.5	25.8	0.46	58.1	5.93	2,659
	1,959,143	1.8	2.2	0.50	5.5	1.44	
Top 10	2,141,599	2.90	26.24	0.59	51.33	3.82	22,505
	1,522,623	2.01	2.79	0.49	6.03	1.76	
Other majors at Los Andes	2,482,154	2.66	25.8	0.55	57.6	5.87	6,069
	1,695,091	1.99	2.2	0.50	5.4	1.53	

Note: Top rows show means and bottom standard deviation. * Based on a clasification over 9 categories of income. Data from all students who started college between 2002 and 2007, and graduated after 2004. The top 10 universities were chosen using SABER PRO scores for schools of at least 1,000 students. Source: Ministry of Education, Colombia.

Table 3a: Baseline results. Effect of the reform on wages.

Economics						
Dep var: Ln wage	(1)	(2)	(3)	(4)	(5)	(6)
Post*Andes	-0.163** [0.0500]	-0.161** [0.0501]	-0.167*** [0.0505]	-0.164** [0.0505]	-0.164** [0.0501]	-0.161** [0.0501]
Post	0.0817** [0.0293]	0.0819** [0.0292]	0.0721* [0.0311]	0.0744* [0.0310]	0.0810* [0.0366]	0.0865* [0.0360]
Andes	0.312*** [0.0304]	0.301*** [0.0301]	0.312*** [0.0304]	0.300*** [0.0301]	0.311*** [0.0304]	0.300*** [0.0301]
Experience	0.135*** [0.00842]	0.154*** [0.0173]	0.137*** [0.00841]	0.154*** [0.0173]	0.135*** [0.0127]	0.156*** [0.0188]
Experience sq		-0.00424 [0.00431]		-0.004 [0.00429]		-0.00429 [0.00431]
Female		-0.0912*** [0.0223]		-0.0908*** [0.0223]		-0.0914*** [0.0224]
Constant	14.16*** [0.0197]	14.20*** [0.0238]	14.13*** [0.0495]	14.17*** [0.0511]	13.96*** [0.0846]	14.19*** [0.0383]
Cohort control	N	N	Y	Y	N	N
Year D	N	N	N	N	Y	Y
Clusters	11	11	11	11	11	11
Obs	3,621	3,621	3,621	3,621	3,621	3,621
R-sq	0.157	0.165	0.157	0.165	0.159	0.167

Standard errors clustered at the individual level.

Control group: students from economics at top 10 schools.

Cohort control: Semiannual GDP growth. Cohort refer to the semester and year the students started school. Year refers to the year of the wage observation.

Ln wage is the natural logarithm of the average monthly wage. Post is a dummy equal to one after the reform, Andes is a dummy equal to one if the student went to Los Andes. Experience is measured in years.

Standard errors in brackets below the coefficients.

*p<0.1, **p<0.05, ***p<0.01

Source: Ministry of Education OLE and SPADIES.

Table 3b: Baseline results. Effect of the reform on wages.

Business						
Dep var: Ln wage	(1)	(2)	(3)	(4)	(5)	(6)
Post*Andes	-0.136*** [0.0410]	-0.136*** [0.0413]	-0.141*** [0.0412]	-0.141*** [0.0414]	-0.135** [0.0412]	-0.136** [0.0414]
Post	0.0952*** [0.0153]	0.0940*** [0.0152]	0.0555** [0.0185]	0.0558** [0.0185]	0.0971*** [0.0189]	0.0991*** [0.0188]
Andes	0.429*** [0.0312]	0.423*** [0.0316]	0.432*** [0.0312]	0.427*** [0.0316]	0.429*** [0.0312]	0.423*** [0.0315]
Experience	0.124*** [0.00517]	0.145*** [0.0115]	0.128*** [0.00512]	0.147*** [0.0115]	0.125*** [0.00782]	0.151*** [0.0120]
Experience sq		-0.00525 [0.00303]		-0.00481 [0.00303]		-0.00635* [0.00302]
Female		-0.0976*** [0.0147]		-0.0969*** [0.0147]		-0.0979*** [0.0148]
Constant	14.06*** [0.0129]	14.11*** [0.0160]	13.96*** [0.0317]	14.00*** [0.0337]	14.15*** [0.0968]	14.10*** [0.0243]
Cohort control	N	N	Y	Y	N	N
Year D	N	N	N	N	Y	Y
Clusters	11	11	11	11	11	11
N	10,348	10,348	10,348	10,348	10,348	10,348
R-sq	0.122	0.130	0.124	0.132	0.122	0.131

Standard errors clustered at the individual level.

Control group: students from business at top 10 schools.

Cohort control: Semiannual GDP growth. Cohort refer to the semester and year the students started school. Year refers to the year of the wage observation.

Ln wage is the natural logarithm of the average monthly wage. Post is a dummy equal to one after the reform, Andes is a dummy equal to one if the student went to Los Andes. Experience is measured in years.

Standard errors in brackets below the coefficients.

*p<0.1, **p<0.05, ***p<0.01

Source: Ministry of Education OLE and SPADIES.

Table 4a: Placebo coefficients

Economics						
Diff in Diff coefficient	(1)	(2)	(3)	(4)	(5)	(6)
Andes	-0.163	-0.161	-0.167	-0.164	-0.164	-0.161
Placebo school 1	-0.145	-0.145	-0.150	-0.148	-0.148	-0.146
Placebo school 2	-0.127	-0.104	-0.129	-0.106	-0.123	-0.099
Placebo school 3	-0.041	-0.053	-0.045	-0.045	-0.044	-0.057
Placebo school 4	-0.039	-0.030	-0.032	-0.035	-0.037	-0.031
Placebo school 5	-0.026	-0.024	-0.022	-0.021	-0.027	-0.025
Placebo school 6	0.080	0.071	0.081	0.074	0.078	0.070
Placebo school 7	0.081	0.073	0.087	0.077	0.085	0.075
Placebo school 8	0.106	0.104	0.109	0.107	0.108	0.103
Placebo school 9	0.113	0.105	0.118	0.109	0.111	0.106
Placebo school 10	0.197	0.220	0.200	0.222	0.206	0.230
Cohort control	N	N	Y	Y	N	N
Year D	N	N	N	N	Y	Y
Clusters	11	11	11	11	11	11
Obs	3,621	3,621	3,621	3,621	3,621	3,621

Cohort control: Semiannual GDP growth. Cohort refer to the semester and year the
Ln wage is the natural logarithm of the average monthly wage. Post is a dummy equal to
Standard erros in brackets below the coefficients.

Source: Ministry of Education OLE and SPADIES.

Table 4a: Placebo coefficients

Business						
Diff in Diff coefficient	(1)	(2)	(3)	(4)	(5)	(6)
Andes	-0.136	-0.136	-0.141	-0.141	-0.135	-0.136
Placebo school 1	-0.085	-0.083	-0.088	-0.087	-0.084	-0.082
Placebo school 2	-0.059	-0.055	-0.052	-0.049	-0.058	-0.054
Placebo school 3	-0.045	-0.045	-0.044	-0.043	-0.046	-0.045
Placebo school 4	-0.040	-0.033	-0.038	-0.035	-0.039	-0.032
Placebo school 5	-0.032	-0.033	-0.034	-0.032	-0.029	-0.032
Placebo school 6	-0.025	-0.017	-0.033	-0.024	-0.020	-0.011
Placebo school 7	-0.008	0.001	0.000	0.008	-0.010	-0.001
Placebo school 8	0.017	0.006	0.025	0.014	0.025	0.015
Placebo school 9	0.072	0.075	0.077	0.079	0.074	0.077
Placebo school 10	0.104	0.100	0.113	0.108	0.107	0.103
Cohort control	N	N	Y	Y	N	N
Year D	N	N	N	N	Y	Y
Clusters	11	11	11	11	11	11
Obs	10,352	10,352	10,352	10,352	10,352	10,352

Cohort control: Semiannual GDP growth. Cohort refer to the semester and year the
Ln wage is the natural logarithm of the average monthly wage. Post is a dummy equal to
Standard erros in brackets below the coefficients.

Source: Ministry of Education OLE and SPADIES.

Table 5: Effect of the reform on wages. Alternative treatment group: students already in school
by the time of the reform.

Panel A: Economics						
Dep var: Ln wage	(1)	(2)	(3)	(4)	(5)	(6)
Post*Andes	-0.161*** [0.0534]	-0.159*** [0.0536]	-0.165*** [0.0538]	-0.161*** [0.0540]	-0.161*** [0.0537]	-0.159*** [0.0539]
Post	0.0737** [0.0288]	0.0726** [0.0289]	0.0669** [0.0308]	0.0681** [0.0309]	0.0713** [0.0344]	0.0768** [0.0343]
Andes	0.312*** [0.0304]	0.300*** [0.0301]	0.312*** [0.0304]	0.300*** [0.0301]	0.311*** [0.0304]	0.298*** [0.0301]
Obs	3,485	3,485	3,485	3,485	3,485	3,485
Panel B: Business						
Post*Andes	-0.118*** [0.0420]	-0.118*** [0.0422]	-0.124*** [0.0172]	-0.124*** [0.0423]	-0.117*** [0.0421]	-0.118*** [0.0424]
Post	0.0925*** [0.0157]	0.0913*** [0.0156]	0.0525** [0.0191]	0.0527*** [0.0185]	0.0933*** [0.0191]	0.0955*** [0.0190]
Andes	0.429*** [0.0312]	0.424*** [0.0315]	0.433*** [0.0811]	0.427*** [0.0316]	0.430*** [0.0312]	0.424*** [0.0315]
Obs.	9,979	9,979	9,979	9,979	9,979	9,979

Standard errors clustered at the student level.

(1) experience. (2) experience, experience squared and gender. (3) experience and cohort controls. (4) experience, experience squared, gender and cohort controls. (5) experience and year dummies. (6) experience, experience squared, gender and year

Cohort control: Semiannual GDP growth. Cohort refer to the semester and year the students started school. Year refers to the year of the wage observation.

Ln wage is the natural logarithm of the average monthly wage. Post is a dummy equal to one if a person studied with the new curricular but was enrolled before the change, Andes is a dummy equal to one if the student went to Los Andes. Experience is measured in years.

Standard errors in brackets below the coefficients.

*p<0.1, **p<0.05, ***p<0.01

Source: Ministry of Education.

Table 6: Cap at three years of experience

Panel A: Economics						
Dep var: Ln wage	(1)	(2)	(3)	(4)	(5)	(6)
Post*Andes	-0.166*** [0.0484]	-0.164*** [0.0486]	-0.170*** [0.0487]	-0.167*** [0.0488]	-0.166*** [0.0485]	-0.164*** [0.0487]
Post	0.0846*** [0.0281]	0.0835*** [0.0281]	0.0751** [0.0304]	0.0752** [0.0304]	0.0829** [0.0343]	0.0858** [0.0343]
Andes	0.314*** [0.0283]	0.305*** [0.0282]	0.313*** [0.0283]	0.304*** [0.0282]	0.313*** [0.0283]	0.304*** [0.0282]
Obs	3,314	3,314	3,314	3,314	3,314	3,314
Panel B: Business						
Post*Andes	-0.129*** [0.0402]	-0.128*** [0.0403]	-0.134*** [0.0403]	-0.132*** [0.0404]	-0.129*** [0.0403]	-0.128*** [0.0404]
Post	0.0953*** [0.0152]	0.0951*** [0.0151]	0.0600*** [0.0184]	0.0609*** [0.0183]	0.101*** [0.0187]	0.104*** [0.0187]
Andes	0.422*** [0.0300]	0.415*** [0.0302]	0.425*** [0.0300]	0.419*** [0.0302]	0.421*** [0.0300]	0.415*** [0.0303]
Obs	9,627	9,627	9,627	9,627	9,627	9,627

Standard errors clustered at the student level.

(1) experience. (2) experience, experience squared and gender. (3) experience and cohort controls. (4) experience, experience squared, gender and cohort controls. (5) experience and year dummies. (6) experience, experience squared, gender and year dummies.

Cohort control: Semiannual GDP growth. Cohort refer to the semester and year the students started school. Year refers to the year of the wage observation.

Ln wage is the natural logarithm of the average monthly wage. Post is a dummy equal to one if a person studied with the new curriculum but was enrolled before the change, Andes is a dummy equal to one if the student went to Los Andes. Experience is measured in years.

Standard errors in brackets below the coefficients.

*p<0.1, **p<0.05, ***p<0.01

Source: Ministry of Education.

Table 7: Effect of the reform on the recruitment process

Dependent variable: 1 if hired and 0 if not	
Andes*Post	-0.167** 0.073
Post	-0.049 0.031
Andes	0.163*** 0.058
Constant	0.112*** 0.023
Obs	438
R squared	0.03

Standard errors below the coefficients

Data from the recruitment process for economist positions from 2008 to 2014

Source: Central Bank of Colombia.

* p<0.1, ** p<0.05, *** p<0.01

Table 8: Placebo test 1—Alternative date of the reform

Panel A: Economics						
Dep var: Ln wage	(1)	(2)	(3)	(4)	(5)	(6)
Fake post*Andes	-0.004 [0.0481]	-0.005 [0.0482]	-0.007 [0.0488]	-0.007 [0.0490]	-0.002 [0.0482]	-0.003 [0.0486]
Fake post	0.012 [0.0458]	0.002 [0.0455]	-0.017 [0.0605]	-0.025 [0.0592]	0.018 [0.0498]	0.015 [0.0497]
Andes	0.313*** [0.0357]	0.300*** [0.0366]	0.315*** [0.0366]	0.301*** [0.0375]	0.309*** [0.0365]	0.294*** [0.0375]
Panel B: Business						
Fake post*Andes	0.016 [0.0838]	0.009 [0.0785]	0.017 [0.0915]	0.009 [0.0847]	0.014 [0.0812]	0.006 [0.0758]
Fake post	0.061 [0.0772]	0.061 [0.0747]	-0.057 [0.184]	-0.054 [0.177]	0.080 [0.0821]	0.082 [0.0782]
Andes	0.420*** [0.0640]	0.417*** [0.0593]	0.423*** [0.0681]	0.420*** [0.0618]	0.420*** [0.0612]	0.416*** [0.0567]

Standard errors clustered at the student level.

(1) experience. (2) experience, experience squared and gender. (3) experience and cohort controls. (4) experience, experience squared, gender and cohort controls. (5) experience and year dummies. (6) experience, experience squared, gender and year dummies.

I take only the students that studied under the old curriculum and set the reform date on the middle of the period (2004-1 for econ and 2003-2 for business).

Standard errors in brackets below the coefficients.

*p<0.1, **p<0.05, ***p<0.01

Source: Ministry of Education.

Table 9: Placebo test 2—Reform evaluated using data from law graduates

Dep var: Ln wage	(1)	(2)	(3)	(4)	(5)	(6)
Date of economics reform						
Post*Andes	-0.00952 [0.0525]	-0.00913 [0.0524]	-0.00696 [0.0535]	-0.00657 [0.0536]	-0.00282 [0.0572]	-0.00261 [0.0573]
Date of business reform						
Post*Andes	-0.0238 [0.0341]	-0.023 [0.0342]	-0.0224 [0.0347]	-0.0216 [0.0348]	-0.0103 [0.0379]	-0.00964 [0.0380]
Obs	3,388	3,388	3,388	3,388	3,388	3,388
R-sq	0.12	0.12	0.12	0.12	0.13	0.13
St errors clustered at the student level.						

(1) experience. (2) experience, experience squared and gender. (3) experience and cohort controls. (4) experience, experience squared, gender and cohort controls. (5) experience and year dummies. (6) experience, experience squared, gender and year dummies.

Cohort control: Semiannual GDP growth. Cohort refer to the semester and year the students started school. Year refers to the year of the wage observation.

Ln wage is the natural logarithm of the average monthly wage. Post is a dummy equal to one if a person studied with the new curriculum but was enrolled before the change, Andes is a dummy equal to one if the student went to Los Andes. Experience is measured in years.

Standard errors in brackets below the coefficients.

*p<0.1, **p<0.05, ***p<0.01

Source: Ministry of Education.

Table 10: Robustness checks

Dep variable: Ln wage	Economics (1)	Economics (2)	Business (3)	Business (4)
<i>Panel a: Controlling for age</i>				
Treatment	-0.162*** [0.0510]	-0.158*** [0.0510]	-0.0952** [0.0410]	-0.0950** [0.0412]
<i>Panel b: Without cohorts that graduated at the same time</i>				
Treatment	-0.159*** [0.0552]	-0.154*** [0.0552]	-0.118*** [0.0437]	-0.118*** [0.0439]
<i>Panel c: Taking graduates from Top 3 schools as control (1)</i>				
Treatment	-0.115** [0.0557]	-0.115** [0.0557]	-0.145*** [0.0472]	-0.145*** [0.0472]
<i>Panel d: Including in the control group only students that could have attended Los Andes</i>				
Treatment	-0.186*** [0.0434]	-0.184*** [0.0441]	-0.152** [0.0640]	-0.151** [0.0626]
<i>Panel e: Without 2007-1 cohort</i>				
Treatment	-0.152*** [0.0510]	-0.146*** [0.0511]	-0.117*** [0.0418]	-0.118*** [0.0420]
<i>Panel f: Controlling for HS exit scores</i>				
Treatment	-0.185*** [0.0469]	-0.180*** [0.0472]	-0.161* [0.0624]	-0.160** [0.0611]
Experience	Y	Y	Y	Y
Experience squared	N	Y	N	Y
Gender	N	Y	N	Y
Cohort effects	Y	Y	Y	Y

Standard errors clustered by individual.

(1) Top 3 schools are Nacional, Javeriana and Rosario.

Standard errors in brackets below the coefficients.

*p<0.1, **p<0.05, ***p<0.01

Source: Ministry of Education.

Table 11: Synthetic control

Dep variable: Ln wage	(1)	(2)
Panel a: Economics		
<i>Control: Industrial Engineering - Javeriana (70.8%)</i>		
Treatment	-0.133** [0.0632]	-0.134** [0.0635]
<i>Control: Industrial Engineering-Nacional (16.3%)</i>		
Treatment	-0.0719 [0.0695]	-0.07 [0.0695]
<i>Control: Oil Engineering-Nacional (7%)</i>		
Treatment	-0.11 [0.0786]	-0.111 [0.0791]
<i>Control: Industrial Engineering- U Norte (6%)</i>		
Treatment	-0.134** [0.0615]	-0.133** [0.0614]
Panel b: Business		
<i>Control: Oil Engineering-Nacional (46%)</i>		
Treatment	-0.197*** [0.0539]	-0.201*** [0.0539]
<i>Control: Business - EAFIT (38.3%)</i>		
Treatment	-0.101* [0.0578]	-0.101* [0.0578]
<i>Control: Industrial Engineering - Javeriana (14%)</i>		
Treatment	-0.0971* [0.0551]	-0.0961* [0.0549]
<i>Control: Law - Andes (1%)</i>		
Treatment	-0.0508 [0.0579]	-0.0506 [0.0577]

Standard errors clustered by individual.

Standard errors in brackets below the coefficients.

The number in parenthesis is the optimal weight.

Column 1 includes experience and cohort controls, column 2 adds experience square and gender.

* p<0.1, ** p<0.05, *** p<0.01

Source: Ministry of Education.

Chapter 2

The Cost of Bad Parents: Evidence from the Effects of Parental Incarceration on Children's Education

Abstract

This paper provides evidence that parental incarceration increases children's educational attainment. I collect criminal records for 90,000 low-income parents who have been convicted of a crime in Colombia, and combine it with administrative data on the educational attainment of their children. I exploit exogenous variation in parental incarceration resulting from the random assignment of defendants to judges with different propensities to convict and incarcerate. My identification strategy differs from the usual judge IV application because I model incarceration as two stage decision problem: First conviction, and then incarceration. I exploit judge leniency along these two different margins. Intuitively, I take advantage of the fact that I can compare children of parents who faced similar judge conviction leniency, but had different incarceration leniency. I derive a new expression that extends the Local Average Treatment Effect concept to a setting with two sources of unobserved treatment heterogeneity. I find that conditional on conviction, parental incarceration increases education by 0.7 years for children whose parents are on the margin of incarceration. This positive effect is larger for boys, violent

crimes, and cases in which the incarcerated parent is the mother.

2.1 Introduction

Over one million children in EU countries, and 2.7 million children in the U.S., have a parent in prison (Sykes and Pettit, 2014).¹ Family environments during the early years, and especially parenting, are major determinants of human development (Heckman, 2013 and Almond et al., 2019), yet there is only a small literature investigating the effects of parental incarceration on children's outcomes. A large body of correlation-based evidence finds negative associations between parental incarceration and a host of important variables such as mental health, education, and crime (Wakefield, 2015). However, households with incarcerated parents are disadvantaged along many dimensions.² Therefore, differences across outcomes means would lead to negatively biased estimates.

In this paper, I estimate the causal effects of parental incarceration on children's educational attainment in Colombia. I exploit exogenous variation in parental incarceration resulting from the random assignment of defendants to judges with different propensities to convict and incarcerate defendants. I construct a new dataset that links sociodemographic data on households with children from SISBEN, Colombia's census of the low-income population, to criminal records for parents scraped from the internet. I find criminal records for approximately 90,000 parents for the years 2005 to 2016. Then, I link the educational outcomes of criminals' children using administrative data on public school enrollment. Finally I web-scrape the children's criminal records after they turn eighteen years old.

I estimate that on average, conditional on conviction, parental incarceration increases education

¹Sykes and Pettit (2014) also estimate that for the U.S., 62% of black children born to high school dropouts will experience the imprisonment of a parent by age 17.

²Even prior to the incarceration event, these households are more likely to be poor and to experience domestic violence (Arditti, 2005; Arditti et al., 2012). In the US, Mumola (2000) finds that 60% of parents in prison reported that they used drugs in the month before their offense, 25% reported a history of alcohol dependence, and about 14% reported a mental illness. Western (2018) also documents that around 60% of parents in prison had experienced childhood trauma, such as domestic violence and sexual abuse.

by 0.7 years for children whose parents were on the margin of going to prison. With an average schooling of 6.8 years, this corresponds to an increase of 10.2%. Under strong monotonicity assumptions, the marginal treatment effect estimates suggests that the benefit of parental incarceration is larger for children of parents who were incarcerated by more lenient judges. Intuitively, on average those who are incarcerated by lenient judges have worse unobserved characteristics, than those incarcerated by the most strict judges. In terms of observed heterogeneity, point estimates suggest that the benefit of parental incarceration is larger when the child is a boy, incarceration was for a violent crime, or the incarcerated parent is the mother, though only the difference in the treatment effects by gender of the child is statistically significant. I also find a U-shaped pattern in the age of the child at the time of the parent's incarceration. Larger positive effects are estimated between ages 0 to 5 and 10 to 15, compared to 5 to 10.

My findings suggest that, on average, parents who are on the margin of incarceration in Colombia are likely to reduce their child's educational attainment if they remain in the household. Research shows that removing a violent parent or negative role model from the household can create a safer environment for a child (Jaffee et al., 2003; Johnson, 2009). Criminal parents may also deplete economic resources in the household. The economic contribution of defendants is likely to be small; Mueller-Smith (2015) finds that in the US, only one-third to two-fifths of incarcerated parents were employed before being charged. Parental incarceration may also reduce the intergenerational transmission of violence, substance abuse, and crime.³ Lastly, parental incarceration may result in the child being placed with an alternative caregiver who has better resources to care for the child. Indeed, I find suggestive evidence which indicates that after an episode of parental incarceration, children often move in with their grandparents. Children are also more likely to move to a household not in SISBEN, which suggests an improvement in economic conditions.

Previous papers in this literature use the random assignment of defendants to judges and their systematic differences in leniency to estimate the causal effects of incarceration on various out-

³For example, using data from Sweden, Hjalmarsson and Lindquist (2007) report significant father-son correlations in criminal activity that begin to appear between the ages 7 and 12, and are fully established by the son's teen years. This result also relates to findings in other fields that conclude that the positive effects of being raised by one's parents depend on the quality of care that the parents can provide (Jaffee et al., 2003).

comes.⁴ These papers omit the fact that there are two distinct margins on which judges are making decisions. Specifically, judges decide first on conviction, and then, for those convicted, they decide on incarceration. When omitting this distinction, researchers compare incarcerated defendants with those who were not convicted and those who were convicted but not incarcerated. These two distinct groups may have different treatment effects that are of interest to policy makers, and respond to different policy concerns. The conviction margin is related to the question of who is, or, who should be convicted of crime; whereas the incarceration margin is related to the question: What is or what should be the punishment for those who are convicted. In my setting this distinction is particularly important because I only observe defendants who are convicted. I show that under treatment effect heterogeneity, the resulting sample selection in the incarceration stage implies that the estimated treatment effect of incarceration does not have the standard local average treatment effect (LATE) interpretation. I derive an easily interpretable expression of the resulting experimental estimates that extends the LATE concept to a setting with two sources of treatment heterogeneity.

I model this situation in a general framework in which treatment can take three values, and is decided upon crossing two thresholds along distinct unobserved margins of selection. In my case, crossing the first threshold decides conviction, and for those convicted, crossing the second threshold determines incarceration. The three treatment outcomes are i) not convicted, ii) convicted and not incarcerated, and iii) convicted and incarcerated. I use the technology to identify treatment effects in a setting with multiple threshold-crossing rules of Lee and Salanie (2018).⁵ Given an instrument for each of the two decision margins, treatment effects related to the second margin (incarceration) can be identified by fixing the crossing of the first threshold, and then exploiting further instrumental variation on the second margin. In the presence of treatment effect hetero-

⁴For previous papers in the incarceration literature see Kling (2006); Aizer and Doyle (2015); Di Tella and Scharrodsky (2013); Mueller-Smith (2015); Bhuller et al. (2016); and Dobbie et al. (2018a), among others. Bhuller et al. (2016) explicitly discusses the multiple dimensions of sentencing decisions when analyzing possible violation of the exclusion restriction, this is however, a distinct concern. The concern here is that the control group of these studies is has a combination of treatment assignment: Not convicted and convicted but not incarcerated, separating this control group may provide important policy results.

⁵Classical approaches to address sample selection omit discussion of treatment effect heterogeneity (Heckman (1978) and Ahn and Powell (1993)).

geneity along the selection margins I estimate a new object which corresponds to the treatment effect of incarceration for those convicted under a given threshold, and who are at the margin of incarceration. Unconditional treatment effects cannot be identified without further assumptions. This weaker identification result is, however, economically relevant: It allows me to estimate the causal effect of incarceration conditional on conviction under a specific conviction leniency. In this empirical exercise, however, I do not find different treatment effects of incarceration along the conviction margin.

To understand the importance of this object, consider a situation in which DNA evidence or phone location records become available in court to decide on conviction. This may change the size and pool of individuals who are found guilty, and as a result treatment effect of incarceration for the new convicted population may also change. This result is also relevant outside the incarceration context. For example, when estimating the returns to STEM versus non-STEM majors, it is important to compare students who had the same underlying probability of attending college. In the framework of my model, we can think of this situation as first deciding on attending college, and then, conditional on college attendance, choosing STEM or non-STEM majors. The returns to STEM may be a function of the underlying probability of attending college.

Contemporaneous to the writing of this article, three papers exploiting judge leniency as an instrument have provided different results using data from Norway, Sweden and Ohio in the US. Dobbie et al. (2019) and Bhuller et al. (2018) find imprecise null effects of parental incarceration on academic achievement for Sweden and Norway, respectively.⁶ For Cleveland, Ohio, Norris et al. (2019) find null effects in test scores or grade repetition, but find that parental incarceration causes children to live in higher socio-economic status neighborhoods as adults, and decreases the likelihood that a child is incarcerated.

⁶There are many differences between Colombia and Scandinavian countries, some of which may drive these different results. First, the size of the treatment is larger in Colombia, where on average prison sentences are 4.4 years, compared with three and eight months in Sweden and Norway, respectively. A second key difference is the potential size of the effects on schooling before college: In Colombia, 31% of the population between 25 and 34 years old has less than a high school degree, whereas this number is 17% for both Norway and Sweden (OECD, 2016). Finally, Norway and Sweden have very generous welfare programs and better education systems compared to those available in Colombia; these programs help insure disadvantaged children and would also point toward smaller treatment effects in the Scandinavian countries.

These results are somewhat in contrast to the large positive effects I find for Colombia. Such heterogeneity points to the importance of understanding the settings and the population who identify the treatment effect in each context. Two key differences can help reconcile these results: first, given the higher incarceration rate in the US, and the lower crime rates in both the Scandinavian countries and the US compared to Colombia, the parents who are incarcerated at the margin in Colombia are more negatively selected than in the US, Norway, or Sweden, in terms of the severity of the crime, but also in terms of income and education. Second, unlike the other papers, my sample consists only of children who lived with their parent prior to the incarceration episode. In the US, half of the parents were not living with their children at the time of incarceration (Parke and Clarke-Stewart, 2002), and as a result the scope for positive effects from removing a parent is very limited. Consistent with this view, other papers that focus on parents living with their children in the US find results similar to mine. Cho (2009) finds that children in Chicago's public schools whose mothers went to prison instead of jail for less than one week are less likely to experience grade retention. Using an event study design, Billings (2018) finds that incarceration improves end-of-grade exams and behavioral outcomes. He also finds, as I do, larger benefits when the mother is the incarcerated parent.

My paper also contributes to the literature on how parents affect their children's outcomes. This includes a large body of papers on the intergenerational effects of human capital (Black et al., 2005; Oreopoulos et al., 2006), wealth (Black et al., 2015), and welfare receipt (Dahl et al., 2014), among other variables. Specifically, my paper contributes to the literature on household structure and children's outcomes, and shows that living with a parent is not always better for children.⁷ Finlay and Neumark (2010) study whether marriage is good for children, and find that unobserved factors drive the negative relationship between never-married motherhood and child education.⁸ In addition, there is mixed evidence on the effects of removing children from their parents and placing

⁷Also see Lang and Zagorsky (2001) who find little evidence that a parent's presence during childhood affects economic well being in adulthood.

⁸There is also a literature in sociology on the effects of marital conflict and divorce on children's well-being. Using longitudinal data, Amato et al. (1995) find that in high-conflict families, children have higher levels of well-being as young adults if their parents divorce rather than stay together.

them in foster care; for South Carolina Roberts (2018) finds positive effects on schooling, Bald et al., (2019) find mixed results across gender and age for Rhode Island, and Doyle (2007, 2008) find negative labor market and crime outcomes for Illinois. My paper contributes to this body of literature with evidence that suggests that children may benefit from the absence of a convicted parent who is at the margin of incarceration.

Finally, my results highlight the importance of parenting, and specifically the costs of bad parents. This calls for a greater governmental role in assisting children from fragile households. Interventions that offer after-school activities can mitigate these costs. Early childhood interventions have been remarkably successful in complementing parental care in very disadvantaged populations (Heckman et al, 2010). These programs can be a starting point to improve the parenting environment of this population.

The rest of the paper is structured as follows. Section 2 provides background on the judicial system in Colombia, and Section 3 describes the data sources and provides summary statistics. Section 4 describes a model to identify causal effects in my setup, Section 5 presents my estimation and results, and Section 6 discusses the results, the mechanism and external validity. Section 7 concludes.

2.2 Background: The Colombian Court System

In this section, I describe the criminal justice system in Colombia: how defendants are processed, how cases are assigned to judges, the types of crimes involved, and the stages of a standard trial.

Figure 2.1 illustrates how defendants are processed in Colombia's criminal justice system.⁹ A criminal record is created when an arrest is made. Once this happens, the police and a randomly assigned prosecutor must present the evidence that motivated the arrest in front of a judge within 36 hours. This judge, who is randomly assigned from the lowest tier of the judicial hierarchy, determines whether the arrest was legal and whether the defendant should await trial in prison.¹⁰

⁹Acuerdo CSJ, 3329.

¹⁰A defendant will go to prison before trial when at least one of the following conditions holds: i) the defendant is

Next, the case is randomly assigned to another judge who will preside over the trial—this is the judge who provides the exogenous variation in conviction and incarceration I use in this paper. In practice, once the first judge decides to continue with the prosecution of a defendant, the case is entered immediately into a software program that assigns a judge at random among the judges in the judicial district and at the court level that the case is designated to; I refer to the district/court level as the “randomization unit.”

Colombia is divided into 33 judicial districts. In the largest cities, a district usually encompasses the city’s metropolitan area, and for the rest of the country, it usually corresponds to a state. Depending on the severity of the charge(s), a case will be randomized within one out of three possible court levels within the judicial district in which the crime was committed. The first level, municipal courts, receive simple cases, such as misdemeanors, property crimes involving small amounts, and simple assault cases. These cases account for 38% of the data. More severe crimes, such as violent crimes, drug- or gun-related crimes, and large property crimes are sent to circuit courts (56%). Lastly, the most severe types of crime, such as aggravated homicide or terrorism, are assigned to a specialized judge (6%).¹¹ On average, there are 20 judges per randomization unit, and the largest district—Bogota—has 55 judges.

Once the judge is assigned, the prosecutor and defense present their arguments to the judge over the course of multiple hearings. The purpose of the first hearing is to formally press charges. In a second hearing, prosecution and defense present all relevant evidence. Next, based on the strength of the evidence, on a third hearing the judge decides on conviction. If the defendant is found guilty, the judge holds a final hearing to determine sentence length and incarceration considering the severity of the crime, potential future harm to society and any aggravating or mitigating factors. The Colombian Penal Code establishes minimum and maximum sentences for each crime, but there is significant discretion on the part of the judge. The general sentencing guidelines range is often quite broad. For example, prison time for possession of 100 grams of cocaine is between

a danger to society, ii) the defendant can interfere with the judicial investigation, or iii) there is reason to believe that the defendant will not appear in court for trial. Art 308. Criminal Proceedings Code.

¹¹ Art 35-37, Criminal Proceedings Code.

five and nine years (Penal Code, Art 376). The judge also determines the crime and severity of the charge the defendant will ultimately be sentenced for—for example, murder versus involuntary manslaughter.

The decision to send a defendant to prison is determined by the length of the sentence. To deal with prison overcrowding, those convicted only serve time in prison when the sentence is longer than a certain threshold.¹² This threshold is set at the national level and has increased overtime. Currently, a sentence equal to four years or less is not served in prison.¹³ As a result, the population that faces a trial is divided into three groups: i) not convicted; ii) convicted and not incarcerated; and iii) convicted and incarcerated. The fact that a portion of the convicted population does not serve time in prison is not a special feature of the Colombian penal system; for example, it is comparable to a sentence of probation in the US.

In Colombia, judges are selected based on their performance on an exam from an open call of attorneys, with specific legal experience requirements for each category of judge. Appointments do not have term limits, and it is common that, over time, judges rise within the judicial hierarchy. The average tenure of a judge is six years, and on average, a judge presides over 344 cases.

While in prison, inmates can receive visits from adults once a week and from their children once a month. The government does not provide special welfare assistance to inmates' families. Unlike in the US, being convicted of a crime does not change one's eligibility for welfare benefits, and in the labor market, it is not common practice to ask about previous convictions, although this information is available online.

¹²This feature is not unique to the Colombian setting(e.g. Italy) and can also be compared to a probation sentence.

¹³In these cases, the only consequence of being convicted is that for the duration of the sentence, the judge must be notified of any change of address or if the convict plans to travel outside the country. Art 63 Penal Code, and Ley 1709 de 2014.

2.3 Data Construction

2.3.1 Data sources

I collect data from several sources. First, I use two waves of Colombia's census of potential beneficiaries of welfare (SISBEN). These data are collected by the government to characterize the country's poor population and to target social programs to them. SISBEN has information on national identification numbers (NINs), household structure, age, gender, education, labor force participation of each household member, and a large set of variables on characteristics and assets of each house (e.g., refrigerator, stove, and floor material, among others). With this information, the government creates a score for each household that summarizes its level of wealth. The score is used to determine eligibility for most public programs—for example, free health insurance, conditional cash transfers, nutrition programs, subsidized housing, and college loans, among many others (Bottia et al., 2012). The first wave, conducted from 2003 to 2005, has data on 31.9 million citizens; the second wave, conducted from 2008 to 2010, has data on 25.6 million citizens.

From this database, I obtain two key elements for my analysis. First, I observe parent and child links when they live in the same household. Second, I use parents' NINs to scrape criminal records that are public and available online. Anecdotal evidence for Colombia suggests that a large share of children with an incarcerated parent were not living with the parent at the time of the crime. My target population is, however, likely to be the most affected by parental incarceration.¹⁴

In Colombia, criminal records from defendants who are convicted are public and available online for 17 out of 33 judicial districts. These 17 districts represent 67% of the population, 69% of homicides, and 83% of property crimes; they include the largest cities in the country; and they are richer and more urban than the 16 districts without data online.¹⁵ Each criminal record includes the name and NIN of the defendant, crime, date of crime, sentence information, and the court type

¹⁴Given how my parent-to-child links are constructed, I focus on parents who are living with the children rather than the biological parents. This definition includes stepchildren when the parent identifies the child as his or her child instead of describing themselves as not being related to the child.

¹⁵The universe of judicial sentences is public; however, they are only available in the nation's National Archives. Criminal records for Bogotá can be found at the following link:
<http://procesos.ramajudicial.gov.co/jepms/bogotajepms/conectar.asp>

and number that handled the case. I collected data on court directories and court identifiers to link each record to a specific judge. There is only one judge per courtroom but judges change over time, I construct the tenure of each judge at each courtroom to assign cases to judges.

I complement these data with individual-level, anonymized records from the Attorney General's Office. This database has information on the universe of criminal cases (including cases that did not result in a conviction), along with courtroom identifiers, date of trial, final verdict, and gender and age of the defendant. I use this information to construct a measure of conviction stringency at the judge level. Finally, I use administrative records of public school enrollment for 2005-2016 with names and NINs to construct a measure of educational attainment. Children's educational attainment is capped at 11, which is the last year of high school in Colombia.

2.3.2 Sample

To construct my sample, I proceed as follows: From SISBEN, I take the NINs of all parents living with their children in the 17 districts that have information online and web-scrape their criminal records. This adds up to 17 million adults. For computational reasons, I only search for records in the district where the person was living at the time of the SISBEN survey. To assess the number of records I miss due to this restriction, I take a 5% random sample and look for their criminal records in all 17 districts. From this, I estimate that I miss 8.6% of the sample due to crimes committed in districts different from the one found in SISBEN. My sample, therefore, includes only poor parents who, at the time of the SISBEN survey, lived with their children, lived in the largest districts of the country, and committed crimes in the district in which they were living.

I find 328,579 criminal records for 256,108 individuals, of which 63,654 have missing fields in at least one of the key variables, such as court identifier, crime, year, or sentence. Half of these records with missing data correspond to Medellin, which is the second largest district after Bogota, and has missing court identifiers in all of their records. I keep only crimes committed after 2005 and after the year of the first SISBEN year records, which results in 193,520 records.¹⁶ Next, I

¹⁶In 2005, there was a reform in the judicial system that renders the two periods incomparable. In the previous

drop all records from court levels for which there was only one judge (5,963 cases dropped), and also in cases in which the number of records per judge in a year is fewer than 15 (44,806). I also only keep courtrooms for which I have judge/year conviction rates from the Attorney General's Office database. This leaves me with 128,792 criminal records from 105,133 adults. I retain only the first conviction in my sample, and collect data on the crime, courtroom identifier, and decisions regarding sentence and incarceration.¹⁷ I merge the criminal records back into the SISBEN data and keep only the first parental conviction in the household. My final data set consists of 91,032 convicted parents.

I link these data to two outcome variables for these children: educational attainment and criminal records. I find school records for 77% of them, similar to the share of children between ages 12 and 17 who attend school (76%, 2005 Census). I also search for criminal records for all children of convicted parents who were 18 years of age by 2017. My final data set consists of 52,419 children born between 1992 and 2007 who have a convicted parent. In the following section, I characterize the population of convicted and incarcerated individuals, as well as their households and children.

2.3.3 Summary statistics

The population in my sample is negatively selected along three margins: education, income and criminal activity. In Table 1, I present socioeconomic characteristics for adults in the overall population, for parents in SISBEN with and without a conviction, and for parents with a conviction, by incarceration status. By comparing column 1 and columns 2 and 3, we see that parents in the SISBEN have fewer years of education, are less likely to have a high school degree, live in larger households, and are more likely to be single than all adults. Among parents in the SISBEN, individuals with a conviction are also negatively selected across a host of variables (column 3 relative to column 2). Convicted adults have fewer years of schooling, are less likely to have a high school

system, a judge served as both prosecutor and judge at the same time, and he or she was anonymous to the defendant. Additionally, at the time of this reform, there were other changes put in place regarding sentencing guidelines.

¹⁷I only keep the first parental conviction to be able to assign the child a unique conviction/incarceration and leniency value.

degree or more (23% vs. 31%), and have lower income scores. They also live in larger households and are more likely to be single (41% vs. 35%, respectively). Adults with criminal records are disproportionately male (84%), they are more likely to work and to be the head of the household than those without a criminal record.¹⁸

Among convicted parents, incarcerated parents have lower education and lower income levels (columns 4 and 5). Gender differences in the probability of incarceration conditional on conviction are far smaller than those in conviction. Incarceration is associated with lower probabilities of working, as well as being the head of the household. Table 2.2 splits the sample by gender. On average, convicted women have lower levels of education relative to convicted men, and they tend to come from poorer households. Compared to men, women are less likely to be the head of the household; yet they are still much more likely to be the heads of their respective households than in the country's overall female population (36% vs. 29%, respectively). Convicted women are also more likely to be single.

Property crimes are the most common type of offense (25%), followed closely by drug-trafficking crimes (24%). Violent crimes account for 20% of the records, followed by gun-trafficking and misdemeanor offenses at 18% and 12%, respectively. Incarceration rates vary substantially by crime. Figure 2.2 ranks crimes by their incarceration rates for selected crimes. Serious crimes, such as kidnapping or rape, have the highest incarceration rates, whereas failure to pay child support, simple assault, and property damage have the lowest. In the middle of the distribution, we find crimes such as drug trafficking, domestic violence, counterfeit currency trafficking, theft, and smuggling, among others.

2.4 Identification

Children from households with incarcerated parents are disadvantaged along many dimensions. As a result, simple comparisons of outcomes of children with and without incarcerated parents

¹⁸In the US context, for example, 29% of parents in state prisons have a high school degree or more, 48% are single, 92% are male, and the median age is 32 (Mumola, 2000).

would lead to negatively biased estimates of the effects of parental incarceration. A common way to address this endogeneity concern is to exploit the random assignment of defendants to judges who differ in their leniency to incarcerate.¹⁹ The intuition of this identification approach is that for a group of defendants on the margin, their incarceration decision will only be determined by whether they were assigned to a harsh or lenient judge.

In this literature, authors have data on the pool of cases randomly assigned across judges and use this to construct their incarceration instrument, as the share incarcerated by each judge —the leave-out mean. They compare incarcerated defendants with non-incarcerated defendants, which includes those who were not convicted as well as those who were convicted but did not receive a prison sentence. These two stages, however, correspond to different policy margins. Conviction is about prosecution and criminal investigation efforts, and incarceration on the other hand, is a matter of punishment or rehabilitation.

In my setting, I only observe defendants who are convicted and conviction is determined after random assignment to a judge, so the observed sample of convicted defendants is not balanced across judges. To address this challenge, I provide a new identification result for a setting in which sample selection invalidates the exogeneity of an instrument. Using the technology in Lee and Salanie (2018), my result extends the insight in Ahn and Powell (1993) to a setting with heterogeneous treatment effects, where I also relax the assumptions on scalar unobservables, linearity in the choice equation, and separability in the outcomes equation. This approach has the additional advantage of estimating a treatment effect of incarceration that has a closer link to policy makers concerns than the one previously estimated in the literature. Instead of comparing incarceration to conviction without incarceration and to those who were found innocent, I only compare incarcerated to the former. In the following section I provide intuition for the identification, after which I formalize this result.

¹⁹See Kling (2006); Aizer and Doyle (2015); Di Tella and Schargrodsy (2013); Mueller-Smith (2015); and Bhuller et al. (2016), among others.

2.4.1 A simplified framework

To fix ideas, let us consider the following framework: Judges are randomly assigned to defendants to make conviction and incarceration decisions by evaluating two distinct attributes of the defendant. When deciding on conviction C , a judge assesses the strength of the evidence of the case at hand. Without loss of generality, the distribution of the strength of the evidence across defendants U^c is uniform $[0,1]$, where zero is the smoking gun and one is no evidence against the defendant. The judge can be one of two types in conviction: harsh (H_c) or lenient (L_c). Harsh judges do not require much evidence to convict a defendant. They have a threshold of 0.8, and thus they convict 80% of defendants; this corresponds to all defendants with a level of evidence below 0.8. Lenient judges require more evidence to convict a defendant, choosing a threshold of 0.2, such that they convict only 20%.

Next, if a defendant is convicted, the judge decides on incarceration I . The judge makes this decision based on an assessment of how harmful the convicted defendant may be to society, and how much punishment the defendant deserves. This trait, which I denote U^I , is also distributed uniformly $[0,1]$. Very harmful defendants have low values of U^I , and non-harmful defendants have values close to 1. Again, regarding incarceration a judge can be either lenient or harsh. A harsh judge (H_I) would send 70% of convicted defendants to prison, whereas a lenient one (L_I) would only incarcerate 30%. It is the same judge making both decisions so a judge can be of one of four types. Figure 2.3 illustrates this situation. The x-axis traces the strength of the evidence the conviction decision is based on. That is, we can order defendants along one relevant dimension—namely, the strength of the evidence in the $[0,1]$ interval. A judge splits the space into two when she or he sets her or his conviction rate: Defendants to the right are free, and defendants to the left are convicted. Similarly, the y-axis traces the defendant's punishment level, which is related to the assessment of predicted future criminal activity; unobserved—to the econometrician, not the judge—crime severity; and any mitigating/aggravating factors or family ties.²⁰ I refer

²⁰As mentioned above, sentencing laws guide the judge's incarceration decisions; however, there is large scope for discretion, even within a specific crime. What this dimension tries to capture are the factors that cause a judge to make different incarceration decisions for criminals who have the same charges.

to this dimension as a measure of the defendants' overall quality. For a fixed level of evidence required for conviction, a judge's incarceration level splits the space of convicted individuals into two: A defendant below the threshold will go to prison, and a defendant above will not.

Due to randomization, all judges start with a statistically identical pool of defendants. However, after the conviction decision is made, the pool of convicted defendants is no longer comparable across judges with different conviction thresholds. Defendants convicted under a judge who requires solid evidence to convict will have, on average, a stronger case against them than those convicted under a judge who convicts even under weak evidence of guilt.

Defendants convicted under a harsh judge can face two types of judges $[H_c, H_I]$ or $[H_c, L_I]$, where the first term refers to the judge's conviction stringency, and the second refers to the incarceration stringency. Similarly, those convicted under lenient judges can also have judges of types $[L_c, H_I]$ and $[L_c, L_I]$. Within these partitions, defendants are balanced across judges: first, because they were randomly assigned to their judge, and second, because they were selected into conviction under the same threshold. As a result, within partitions, there is exogenous variation in the probability of going to prison. For example, convicted defendants who were assigned to a $[H_c, L_I]$ judge face a 30% chance of incarceration, whereas those assigned to a $[H_c, H_I]$ judge face a 70% probability. Figure 2.4 illustrates this argument. This means that for 40% of defendants whose harmfulness assessment is located above the worst 30% of the population, but still in the bottom 70%, incarceration is only a function of judge assignment. Thus, I will be able to estimate LATE-type parameters for defendants who fall into this range.

Specifically, for this example I estimate the following two LATE parameters:

$$LATE_{H_c} = E[Y(t_I) - Y(t_c) | U^c < 0.8, 0.3 < U^I < 0.7]$$

and,

$$LATE_{L_c} = E[Y(t_I) - Y(t_c) | U^c < 0.2, 0.3 < U^I < 0.7]$$

Where $LATE_{H_c}$ is the causal effect of incarceration relative to conviction for those convicted under a harsh judge ($U^c < 0.8$), and $LATE_{L_c}$ is the one for conviction under a lenient judge. $Y(t_l)$ and $Y(t_c)$ represent counterfactual outcomes (years of education of the child) for incarceration (I) and conviction (C), and U^c traces the selection on the conviction stage.

$$LATE_{H_c} = \frac{E[Y|H_c, H_l] - E[Y|H_c, L_l]}{E[T = I|H_c, H_l] - E[T = I|H_c, L_l]}$$

Where $T = I$ in the denominator represents treatment assignment equal to incarceration. Similarly, we can have the analogous expression for $LATE_{L_c}$.

2.5 Estimation

To apply the identification result of the previous section, I start by estimating the sample analogs of $P_c(Z)$ and $P_l^*(Z)$ in the model. The interpretation of these variables is the probability of being convicted/incarcerated, given the assignment to a specific judge. Following the literature, these are estimated as judge fixed effects from regressions after parsing out variation at the unit at which the randomization of judges occurred and specific case characteristics. That is, the conviction/incarceration decision can be decomposed into a portion that is related to the individual, the judge, the offense, and the randomization unit/year. I do this as follows:

$$D_{itorz} = \gamma_{rt} + \gamma_o + \varepsilon_{itorz}$$

Where D_{itorz} corresponds to a conviction or incarceration dummy, i indexes individuals, t year, o offense, r court-level/judicial district and z judge. γ_{rt} corresponds to randomization-level fixed effects, which is a court-level/judicial-district by year-level fixed effect. γ_o is a offense-level fixed effect (161 different crimes); and ε_{itorz} is a mean zero term. Following the literature, I estimate the judge instrument $\widehat{p_{z-i}}$ for defendant i to be the following leave-out estimator:

$$\widehat{p_{zi}} = \frac{1}{n_z - 1} \sum_{k \neq i} \widehat{res_{z,k}}$$

where n_z is the number of cases of judge z , and res_{zk} is the residual from a regression of the conviction/incarceration dummy on γ_{rt} and γ_o .

Figure 2.5 shows the distribution of D_{itor} at the judge level, and $\widehat{p_z}$ for both conviction and incarceration. From the graph, we can see that although court-level/year and crime-level fixed effects explain most of the variation, judge's fixed effects still represent a sizable share of the variance in conviction and incarceration.

2.5.1 Instrument validity

Next, I examine how much judge fixed effects predict individual-level decisions by estimating a first-stage regression, as follows:

$$D_{itorz} = \beta_0 + \widehat{p_{zi}} + \beta_1 X_i + \varepsilon_{itorz}$$

As before, D_{itorz} corresponds to the conviction or incarceration dummy, and p_z is the leave-out mean of judge z assigned to person i . I run this regression with and without controls X_i . In the conviction regression, where I use anonymized data from the Attorney General's Office, I can only control for age, gender, and number of crimes charged. In the incarceration regression, I control for schooling, income, occupation, gender, year of birth, and year in the survey. According to the results in Table 2.3, judges have a strong influence on conviction and incarceration decisions. The estimates are highly significant and suggest that being assigned to a judge with a 10 percentage point higher conviction/incarceration rate increases the defendant's probability of conviction and incarceration by seven and eight percentage points, respectively. This relationship is robust to the inclusion of controls, as expected by random assignment. Figure 2.6 depicts this first-stage relationship for conviction (left panel) and incarceration (right panel). These graphs show a strong positive relationship between the instrument and individual trial decisions. The F-stats on the first

stage correspond to regressions on judge dummy variables to account for the true dimensionality of the instruments. These F-stats are above the critical value for the leave-out mean instrument for weak instruments (see Figure 4 in Stock et al., 2002). See Section 5.4 for a further discussion about the F-stat.

Recall from the previous section that the variation in incarceration stringency conditional on a level of conviction stringency is what identifies treatment effects in this context. Figure 2.7 shows a scatter plot of both conviction and incarceration fixed effects. From the graph we can see that there is substantial variation along the incarceration axis for each conviction rate.

For the instrument to be valid, the judge's fixed effects must be orthogonal to the defendant's characteristics. I test this in the anonymized data from the Attorney General's Office, where the universe of cases the judge has heard is available. Table 2.4 checks the balance across defendants for my judge-stringency measures for conviction and incarceration. Across gender, age, and type of crime—which are the only variables available in these data—I find no individual or joint statistical significance. In addition, the identification result is supported by the observation that once P_c is fixed, the pool of convicted defendants is balanced across judges. I test whether covariates are associated with incarceration stringency for the convicted sample, once I split the sample by conviction group (low, medium, or high) or control for the conviction level with a polynomial of P_c . In Table 2.5, I test the individual and joint significance of variables associated with education, income, and occupation status, and find no evidence of a relationship with judge stringency.

To interpret the results of the IV as the causal effect of incarceration, judge stringency must only affect child's outcomes through incarceration. This may not be the case if the judge fixed effects capture other dimensions of trial decisions, such as fines or guilt (Mueller-Smith, 2017). In my setting, this is less of a concern because in the case of Colombia, fines are rare and only associated with large property crimes; and because I model the conviction decision directly.

Finally, I also require that conviction or incarceration decisions made by a lenient judge would also have been made by a stricter judge; this is called the monotonicity assumption. One testable implication of monotonicity is that first-stage estimates should be non-negative for all sub-samples

(Bhuller et al, 2016). That is for example if a judge is lenient, he or she is going to be lenient for both women and men, and for both violent crimes and nonviolent crimes. To test this assumption, I construct judge fixed effects for just one group in the population, for example, for men and use this fixed effect in a first-stage regression to predict individual conviction and incarceration for women. I do this for gender, type of crime, and age group. I find positive first-stage estimates across all slices of the data, which supports the monotonicity assumption. However, if only these weaker monotonicity holds inference is constrained. In particular, it does not allow for the identification of marginal effects along the entire distribution of judge propensities, as can be achieved in the conventional framework. The weaker assumptions rely on averaging across the entire set of judges, while identification of marginal effects throughout the distribution requires assumptions to hold judge by judge (Norris, 2019). In Table 2.6 I test pairwise monotonicity following Norris (2019) and find I can not reject monotonicity across individuals characteristics, and it is only rejected for property vs not property crimes.²¹ Finally, Frandsen et al (2019) show that under the usual assumptions, average outcomes by judge will be a continuous function with bounded slope of judge propensities to incarcerate. Intuitively, if this is not the case, it implies that either judges influence outcomes beyond their propensity to assign treatment, or judges disagree on their implicit ordering of which defendants should be treated. Based on that result, they develop a test that jointly test violations to the monotonicity assumption and the exclusion restriction. In Table 2.7 I implement their test and I find there is no evidence of violation of this assumption.

2.5.2 Results

Following the identification result, I need to account for the different levels of conviction stringency at which defendants were found guilty. I do this in two ways: First, I sort my data by stringency in the conviction stage (P_c) and split the sample into terciles: low ($0.7 < P_c < 0.88$), medium ($0.88 < P_c < 0.9$), and high ($0.9 < P_c < 1$) conviction levels. Second, I pool the data and add a second-degree polynomial on P_c with interaction terms. This last estimate can be interpreted as

²¹I split judge leniency across this characteristic and find very similar point estimates.

an average effect across the different conviction thresholds. The first column of Tables 2.8 and 2.9 show the pooled regression, and the following three columns show the regressions for the split sample.

I begin by showing the OLS estimate of this design. Table 2.8 shows a regression of parental incarceration on years of education. Following Abadie et al. (2017), I cluster standard errors at the randomization level. Without controls, a child whose parent went to prison has around 0.3 fewer years of schooling than a child whose parent did not. Once I add controls, this difference reduces drastically to less than 0.1 years. Still, we expect that incarcerated parents are negatively selected on unobservables that cannot be accounted for, so -0.1 years is a lower bound on the causal effect.

Next, Figure 2.8 shows a graphical representation of the reduced-form regression. This graph plots the distribution of judges' incarceration fixed effects against the predicted years of education from a local polynomial regression. From the graph, we can see that there is a strong positive relationship between judge stringency in incarceration and years of education. That is, as we move to the right, where the probability of having a parent in prison increases exogenously, I estimate that the years of education also increase. The top panel of Table 2.9 shows the regression results for this reduced form: I estimate large increases in years of education for all specifications the increase in years of education is statistically significant. Finally, the bottom panel of Table 2.9 shows results from the IV; I estimate that having an incarcerated parent increases years of schooling by around 0.7 years on average for all convictions levels. These estimates are statistically different from zero. I find that the increase in years of education is mostly accrued through higher graduation rate from middle school. There are positive treatment effects for all grades, but the effect is larger for 9th grade which corresponds to the last grade of middle school.

I also study how parental incarceration affects the chance that the child is later convicted of a crime. For this exercise, I restrict the data to children who were 18 years old by 2017, so that their criminal records would be public. Figure ?? graphically depicts reduced-form estimates of judge stringency on conviction probability; the effect is close to zero. However, the analysis is under-powered to detect to estimate reasonably sized treatment effects. This is not surprising,

since conviction is a low incidence event; only 1.6% of children had a criminal record, and the difference in the OLS is only 0.1 pp.

2.5.3 Heterogeneity

In this section I examine the heterogeneity of the results along observables and unobservables. In my context, marginal treatment effects (MTE) are particularly interesting, because they trace the causal effect of incarceration along parents' unobserved characteristics (U^I) that matter for incarceration and that are correlated with defendants' quality, broadly defined. What this exercise does is to evaluate the possibility of different effects of parental incarceration given the type of defendant that is going to prison, which is characterized by his or her location along the y-axis of Figure 2.3. The intuition is as follows: Parents who are incarcerated under the most lenient judges have worse characteristics than those incarcerated under strict judges. In other words, a strict judge incarcerates almost everyone, but a lenient judge incarcerates only the worst defendants, so that those incarcerated under relatively lenient judges are more negatively selected.²² I follow Heckman and Vitlacil (2005) to estimate this MTE. Under stronger monotonicity assumptions, I find that at the 5% level, there are heterogeneous treatment effects along parental quality (Figure 2.9). Specifically, I find that the positive effects of incarceration on schooling accrue when the worst defendants go to prison.

The magnitude of the effect of parental incarceration on children's education is a function of the relationship between the parent and the child prior to the incarceration episode, the type or quality of this parent, and the role of the child in the household. To document this heterogeneity, I estimate the IV regression for different subgroups in the data. Following previous literature in economics, as well as that in psychology and sociology, I estimate different regressions by gender of the child, gender of the parent, child's age at the time of the incarceration episode, birth order, and the nature of the offense—violent, property, drug- or gun-related, and misdemeanor. In Table

²²I look at this empirically and find that among incarcerated defendants, those incarcerated under stricter judges tend to have fewer and less severe charges. This follows almost directly from the definition of leniency, but also helps to illustrate the way in which these defendants are better.

2.10 I show IV results for the pooled model for these different groups in the data.

According to the estimates, the benefits of parental incarceration are larger for boys than girls, and this difference is statistically significant. Specifically, I find that boys' schooling increases by 0.86 years, whereas girls' schooling increases by 0.36 years. This result is consistent with previous research in psychology and economics, which documents that boys are more vulnerable than girls to negative experiences in the household (Bertrand & Pan (2013); Autor et al. (2016); Parke & Clarke-Stewart (2002); Hetherington et al., 1998). Specifically, Autor et al. find that boys, relative to their sisters, have higher rates of disciplinary problems, lower achievement scores, and fewer high school completions when growing up in disadvantaged environments.

I split the sample by gender of the parent and find that incarceration is more beneficial in cases in which the mother is the one going to prison. This result might be surprising at first glance. However, it is important to bear in mind that children's well-being is more closely affected by their mothers' behavior because of their main role as primary caregivers, and that criminal women are more negatively selected than criminal men (Table 2.2). This result is consistent with the findings of previous research in the US, where Billings (2018) and Turanovic et al. (2012) estimate larger positive effects from maternal incarceration.²³

A source of heterogeneity associated with the quality of the parent going to prison is the type of crime they committed. Thus, in the lower panel of Table 2.10 I split the sample by crime categories: violent, property, drug-related and gun-related. The largest benefits are observed for defendants convicted for violent crimes, whereas the smaller benefits are for property crimes. These differences, however, are not statistically significant. Nonetheless, this is in line with the previous result on unobserved heterogeneity, in which the positive effects are a function of how good the defendant is as a parent.

Lastly, I look at heterogeneous effects depending on the age of the child at the time of parental conviction. I split the sample into three groups: children who were 0 to 5 years, 6 to 10 years, and

²³It is also the case that in the US, incarcerated women have worse socioeconomic backgrounds than incarcerated men (Harrison & Beck, 2006). In addition, Glaze and Maruschak (2008) survey incarcerated parents and find that 60% of imprisoned mothers, compared to 16% of fathers, have histories of being physically or sexually abused.

11 to 15 years at the time of parental conviction. I find a U pattern in the effects on schooling. Studies in developmental psychology conclude that children in the first age group are the most vulnerable, as they do not yet have the abilities and skills to process trauma on their own (Johnston, 1995). These skills and abilities develop over time, and help children cope with distress. On the other hand, the increase in the positive effect in the later years may be the result of how salient the decision is to continue in school or drop out at older ages.

2.5.4 Robustness

In this section I go over various exercises that evaluate the robustness of the results in the paper along different dimensions.

In Table 2.3 I report the first-stage regression on incarceration, and in the bottom of the table I report the F-test on the excluded instruments. This F-test corrects for the fact that the dimensionality of the instrument is the number of judges and not one (my measure of judge leniency). With this correction, the F-stats are low, but above the critical values for weak instruments. The consequence of weak instruments is that 2SLS-IV estimates will be biased toward the OLS (Stock et al., 2002). In my context, given that the OLS estimates are negative, the bias of the OLS is also negative, and the 2SLS IV estimates are positive, this means that we could expect even larger positive effects. To assess the size of this residual bias, I estimate the IV using the LIML estimator, which is less sensitive to weak instruments—the bias does not increase with the number of instruments (Rothenberg, 1993; Stock et al., 2002). I find that the 2SLS and LIML estimator are very close and both are around a point estimate of 0.8 years.

In the Results section, I show my preferred specifications for the estimates of the effect of parental incarceration on educational attainment. This decision to split the sample into three groups of P_c was arbitrary. To assess the robustness of the results, in Figure ?? I instead order observations along P_c , and run multiple regressions on a rolling window of 18,000 observations over P_c , moving the window 500 observations each time. I estimate that for each sample, I find a positive effect of incarceration on education.

Lastly, as a placebo check, I evaluate whether there are differences in schooling for children of incarcerated versus non-incarcerated parents before the date of the sentence. I estimate that there is no supporting evidence that the positive effects I estimate are the result of preexisting differences in educational attainment.

2.6 Mechanisms

2.6.1 What explains the positive effect?

The results presented here suggest that living with a convicted parent has negative consequences for their children. There are many reasons to believe that this is plausible. First, criminals are more likely to exert psychological and physical violence at home, and this can often be detrimental to a child's well-being. In the US context, Western et al. (2004) find that incarcerated men engage in domestic violence at a rate about four times higher than the rest of the population. Furthermore, research in psychology documents that spending time with parents who engage in high levels of antisocial behavior is associated with more conduct problems for their children (Jaffee et al., 2003). This literature concludes that the salutary effects of being raised by married biological parents depend on the quality of care the parents provide.

Second, Chimeli and Soares (2017) document the causal effect of trading illegal commodities on violence. In light of their work, we can expect that households that take part in illegal businesses face constant violence or threats of violence related to guaranteeing property rights or resolving disputes within the business, all of which affect the quality of life in a household. There is also literature on the intergenerational transmission of violence, substance abuse, and crime. Specifically, in the role-model theory, in which children directly observe and model their parents' behavior, incarcerating parents could be beneficial, as it removes bad role models from the house and forces children to update their beliefs about the consequences of criminal behavior (Hjalmarsson and Lindquist, 2012). Beyond intergenerational transmission, childhood exposure to negative behaviors is documented to have direct adverse effects on outcomes in both childhood and adulthood

(Balsa, 2008; Chatterji and Markowitz, 2000), all of which helps explain the positive estimates in this paper.

2.6.2 How does the environment of the child change?

To characterize the changes that households and children experience after an episode of incarceration, I analyze households for which I have two observations in the SISBEN (44% of cases), in which the parent was convicted of a crime between observations. Appearing in both waves of the SISBEN is not random; on the contrary, leaving the sample is associated with an improvement in living standards. This is particularly relevant for children who might be moving to a household outside of SISBEN after the episode of parental incarceration. With this caveat, Table 2.11 shows suggestive evidence that incarceration is associated with an increase in labor force participation (LFP) of the spouse, a worsening of the income score of the household, and a decrease in the probability of a male as the head of the household. I also find that the probability of living with grandparents increases and the probability of being in the second wave of SISBEN decreases, suggesting that incarceration induces children to move in with relatives who are better off financially.

2.6.3 Parents at the margin

To derive policy implications, it is important to acknowledge the local feature of my estimates. This paper estimates the effects of parental incarceration for a particular sub-population: children of convicted poor parents at the margin of incarceration. A large share of those convicted—for example, those guilty of murder or rape—would be incarcerated regardless of judge assignment, and this paper cannot provide any insights into the effects on educational attainment of the children of those individuals. At the other end of the distribution, defendants convicted of minor crimes will also avoid prison, regardless of judge assignment. Defendants convicted of drug- or gun-trafficking, domestic violence, and medium-sized property crimes compose the complier group in my estimation, and they are the group my estimates apply to. This marginal population, however, is particularly relevant because it is the population that is more likely to be affected by policy

interventions to the criminal justice system. Following Dahl et al. (2014), I find that compliers make up approximately 29.8% of the sample.²⁴

2.6.4 External validity and policy implications

To assess the external validity of my results, I provide a framework motivated by my heterogeneity analysis, which links parental quality to parenting treatment effects and the probability of incarceration. Figure 2.10 summarizes this framework. The x-axis traces parental quality; as we move to the right, parental quality increases. The y-axis measures the treatment effect of parenting: Having better parents is better for children. Most importantly, however, there is a segment in the support of parental quality for which parents are detrimental for children. The secondary y-axis measures incarceration probability: In the model, the probability of being incarcerated decreases when parental quality increases. Each society chooses a level of incarceration, which is characterized by a threshold in the support of parental quality. This threshold determines the average effect of incarcerating parents (the gray area in Figure 2.10). To determine how much the results in this paper apply to other settings, we need to think about the location of the incarceration threshold along the parental quality axis and the shape of the function of the treatment effects' of parents in each country. Countries with higher incarceration rates will incarcerate, on average, better parents than those with lower rates, and as a result we should expect lower benefits or even costs from parental incarceration. We can also expect a much flatter function of treatment effects of parenting in generous welfare states, such as the Nordic countries, in which children's education and health vary less with parental characteristics. As a consequence, we would find smaller treatment effects of parental incarceration (both positive and negative). Similarly, some of the estimates in the lit-

²⁴Parental compliers are defendants who would have received a different incarceration decision had their case been assigned to the most lenient judge instead of the strictest judge. We can define the size of this group (π_c) as follows:

$$\pi_c = \text{Prob}(\text{Incarceration} = 1 | z_j = \bar{z}) - \text{Prob}(\text{Incarceration} = 1 | z_j = \underline{z})$$

where \bar{z} and \underline{z} correspond to the incarceration rates of a judge at the 99th and 1st percentiles, respectively. Because of monotonicity, the share of parents who would go to prison regardless of the judge assigned to their case—always takers—is given by the incarceration rate for the most lenient judge and is equal to 22.5%. On the other hand, 47.7% of the sample are children of never takers who would not go to prison no matter which judge was assigned to their case. I estimate that children of compliers make up approximately 29.8% of the sample.

erature (Norris et al 2019 and Dobbie et al 2019) consider birth parents who may not necessarily co-reside with their children, in this framework we can hypothesized that it translates to smaller treatment effect of parents and as a result into a smaller effect of parental incarceration.

2.7 Conclusions

The rise in incarceration has led to an increase in the number of children growing up with a parent in prison. In this paper, I estimate the causal effects of parental incarceration on educational attainment in Colombia. My results suggest that children benefit when their convicted parents are incarcerated. Specifically, I estimate that parental incarceration increases schooling by 0.7 years on average.

I conclude with a discussion of three important limitations of this paper. First, I consider only the short-term effects of parental incarceration. This is important, as these parents eventually leave prison and will perhaps return to live with their children. Further, if incarceration decreases one's human capital and social and emotional skills, the type of parent who returns after incarceration can be much worse than the one who left. In that case, the long-term effects may be very different from what I estimate here. Another significant limitation of this paper is that, effectively, I can only study one outcome variable. As shown by Dobbie et al. (2018), parental incarceration can have sizable effects on other variables such as earnings and teen pregnancy. These are important results that help characterize the complex shock of having an incarcerated parent, but due to data limitations, I cannot explore them here. Finally, my paper only offers suggestive evidence on the mechanisms that explain the positive effects of parental incarceration on children's educational attainment, further research is required to characterize the obstacles children face in these households in order to provide informed policy recommendations.

2.8 References

Abadie, A., Athey, S., Imbens, G.W. & Wooldridge, J., (2017). When Should You Adjust Standard Errors for Clustering? (No. w24003). National Bureau of Economic Research.

Abadie, A. (2003). Semiparametric Instrumental Variable Estimation of Treatment Response Models,” *Journal of Econometrics*, 113(2), 231-263.

Aizer, A. & J. J. Doyle (2015). Juvenile Incarceration, Human Capital and Future Crime: Evidence from Randomly-Assigned Judges. *The Quarterly Journal of Economics* 130 (2), 759–803.

Ahn, H., & Powell, J. L. (1993). Semiparametric estimation of censored selection models with a nonparametric selection mechanism. *Journal of Econometrics*, 58(1-2), 3-29.

Angrist, J. (1995). Conditioning on the probability of selection to control selection bias. NBER Technical Working Paper No. 181, June 1995.

Amato, P. R., Loomis, L. S., & Booth, A. (1995). Parental divorce, marital conflict, and offspring well-being during early adulthood. *Social Forces*, 73(3), 895-915.

Arditti, J.A., 2015. Family process perspective on the heterogeneous effects of maternal incarceration on child wellbeing. *Criminology and Public Policy*, 14(1), pp.169-182.

Arditti, Joyce (2012). Parental incarceration and the family: Psychological and social effects of imprisonment on children, parents, and caregivers. New York, NY: New York University Press.

Arditti, Joyce, Sara A. Smock, & Tiffaney S. Parkman (2005). It’s been hard to be a father: A qualitative exploration of incarcerated fatherhood. *Fathering* 3:267–83.

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2016). Family disadvantage and the gender gap in behavioral and educational outcomes (No. w22267). National Bureau of Economic Research.

Bald, A., Chyn, E., Hastings, J. S., and Machelett, M. (2019). The Causal Impact of Removing Children from Abusive and Neglectful Homes (No. w25419). National Bureau of Economic Research.

Balsa, A. I. (2008). Parental problem-drinking and adult children’s labor market outcomes. *Journal of Human Resources*, 43(2), 454-486.

Bertrand, M. & Pan, J., 2013. The trouble with boys: Social influences and the gender gap in disruptive behavior. *American Economic Journal: Applied Economics*, 5(1), pp.32-64.

Bhuller, Manudeep, Gordon B. Dahl, Katrine V. Løken, & Magne Mogstad. 2017. "Incarceration, Recidivism and Employment." *Quarterly Journal of Economics*. 22648.

Bhuller, M., Dahl, G. B., Loken, K. V., & Mogstad, M. (2018). Intergenerational effects of incarceration. In *AEA Papers and Proceedings* (Vol. 108, pp. 234-40).

Billings, Stephen (2017) Parental Arrest and Incarceration: How Does it Impact the Children? (Preliminary draft)

Black, S.E., Devereux, P.J. & Salvanes, K.G., 2005. Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American Economic Review*, 95(1), pp.437-449.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, 120(2), 669-700.

Chimeli, A. B., and Soares, R. R. (2017). The use of violence in illegal markets: Evidence from mahogany trade in the Brazilian Amazon. *American Economic Journal: Applied Economics*, 9(4), 30-57.

Cho, Rosa M. 2009a. "The Impact of Maternal Imprisonment on Children's Probability of Grade Retention: Results from Chicago Public Schools." *Journal of Urban Economics*, 65(1): 11-23.

Cho, Rosa M. 2009b. "The Impact of Maternal Incarceration on Children's Educational Achievement: Results from Chicago Public Schools." *Journal of Human Resources*, 44(3): 772-797.

Criminal Proceeding Code (2004). *Codigo de Procedimiento Penal*. Ley 906 de 2004; Bogota, Colombia.

Cunha, F., I. T. Elo, and J. Culhane (2013). Eliciting maternal expectations about the technology of cognitive skill formation. Working Paper 19144, NBER.

Currie, J. and Moretti, E., 2003. Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, 118(4), pp.1495-1532.

Dahl, G. B., A. R. Kostøl, and M. Mogstad (2014). Family Welfare Cultures. *The Quarterly Journal of Economics* 129 (4), 1711–1752.

Di Tella, R. and E. Schargrodsky (2013). Criminal Recidivism after Prison and Electronic Monitoring. *Journal of Political Economy* 121 (1), 28–73.

Dobbie, W., Goldin, J., & Yang, C. S. (2018). The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges. *American Economic Review*, 108(2), 201-40.

Dobbie, W., H. Grönqvist, S. Niknami, M. Palme and M. Priksk (2018). The Intergenerational Effects of Parental Incarceration. NBER Working Paper, January.

Doyle Jr, J. J. (2007). Child protection and child outcomes: Measuring the effects of foster care. *American Economic Review*, 97(5), 1583-1610.

Doyle Jr, J. J. (2008). "Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care." *Journal of Political Economy*, 116(4): 746-770.

Ehrensaft, M.K., Cohen, P., Brown, J., Smailes, E., Chen, H. and Johnson, J.G., 2003. Intergenerational transmission of partner violence: a 20-year prospective study. *Journal of consulting and clinical psychology*, 71(4), p.741.

Finlay, K., and Neumark, D. (2010). Is marriage always good for children? Evidence from families affected by incarceration. *Journal of Human Resources*, 45(4), 1046-1088.

Frandsen, B. R., Lefgren, L. J., and Leslie, E. C. (2019). Judging Judge Fixed Effects (No. w25528). National Bureau of Economic Research.

Furstenberg, F. F., Jr. (1995). Fathering in the inner city: Paternal participation and public policy. In W. Marsiglio (Ed.), *Research on men and masculinities series*, 7. *Fatherhood: Contemporary theory, research, and social policy* (pp. 119-147). Thousand Oaks, CA, US: Sage Publications, Inc.

- Fomin, S. V. (1999). Elements of the theory of functions and functional analysis (Vol. 1). Courier Corporation.
- Imbens, G.W., and D. B. Rubin (1997). Estimating Outcome Distributions for Compliers in Instrumental Variables Models,” *The Review of Economic Studies*, 64(4).
- Harrison, P. M. and A. J. Beck (2006). Prison and jail inmates at midyear 2005.
- Hart, B. and T. R. Risley (1995). Meaningful differences in the everyday experience of young American children. Baltimore, MD: P.H. Brookes.
- Heckman, J. J., Urzua, S., and Vytlacil, E. (2006). Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics*, 88(3), 389-432.
- Heckman, James J., and Edward Vytlacil. 2005. “Structural Equations, Treatment Effects, and Econometric Policy Evaluation.” *Econometrica*, 73(3): 669-738.
- Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Q. Yavitz (2010a). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative Economics* 1 (1), 1–46.
- Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Q. Yavitz (2010b). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics* 94 (1–2), 114–128.
- Heckman, J. J. (2013). Giving kids a fair chance. Mit Press.
- Hetherington, E. M., Bridges, M., and Insabella, G. M. (1998). What matters? What does not? Five perspectives on the association between marital transitions and children’s adjustment. *American Psychologist*, 53(2), 167.
- Hjalmarsson, Randi, and Matthew J. Lindquist. 2011. “The Origins of Intergenerational Associations in Crime: Lessons from Swedish Adoption Data.” *Labour Economics*, 20: 68-81.
- Hjalmarsson, Randi, and Matthew J. Lindquist. 2012. “Like Godfather, Like Son: Exploring the Intergenerational Nature of Crime.” *Journal of Human Resources*, 47(2): 550-582.
- Hjalmarsson, Randi, Helena Holmlund, and Matthew J. Lindquist. 2015. “The Effect of Education on Criminal Convictions and Incarceration: Causal Evidence from Micro-data.” *Economic*

Journal, 125(587): 1290-1326.

Jaffee, S. R., Moffitt, T. E., Caspi, A., and Taylor, A. (2003). Life with (or without) father: The benefits of living with two biological parents depend on the father's antisocial behavior. *Child development*, 74(1), 109-126.

Johnson, R. (2009). Ever-increasing levels of parental incarceration and the consequences for children. Do prisons make us safer? The benefits and costs of the prison boom, 177-206.

Kalil, A., 2015. Inequality begins at home: The role of parenting in the diverging destinies of rich and poor children. In *Families in an era of increasing inequality* (pp. 63-82). Springer, Cham.

Kim-Cohen, J., Moffitt, T.E., Taylor, A., Pawlby, S.J. and Caspi, A., 2005. Maternal depression and children's antisocial behavior: nature and nurture effects. *JAMA Archives of general psychiatry*, 62(2), pp.173-181.

Kling, J. R. (2006). Incarceration Length, Employment, and Earnings. *The American Economic Review* 96 (3), 863–876.

Lang, K., and Zagorsky, J. L. (2001). Does growing up with a parent absent really hurt?. *Journal of human Resources*, 253-273.

Lee, Sokbae, and Bernard Salanié. "Identifying effects of multivalued treatments." (2018).

Lefgren, L., Sims, D. and Lindquist, M.J., 2012. Rich dad, smart dad: Decomposing the intergenerational transmission of income. *Journal of Political Economy*, 120(2), pp.268-303.

Lyle, D. S. (2006). Using military deployments and job assignments to estimate the effect of parental absences and household relocations on children's academic achievement. *Journal of Labor Economics*, 24(2), 319-350.

McLanahan, S., Tach, L., and Schneider, D. (2013). The causal effects of father absence. *Annual review of sociology*, 39, 399-427.

Mueller-Smith, M. (2017). *The Criminal and Labor Market Impacts of Incarceration*. University of Michigan Working Paper.

Murray, Joseph, and David P. Farrington. 2005. "Parental Imprisonment: Effects on Boys' Antisocial Behaviour and Delinquency Through the Life-Course." *Journal of Child Psychology*

and Psychiatry, 46(12): 1269-1278.

Murray, Joseph, David P. Farrington, and Ivana Sekol. 2012. "Children's Antisocial Behavior, Mental Health, Drug Use, and Educational Performance After Parental Incarceration: A Systematic Review and Meta-analysis." *Psychological Bulletin*, 138(2): 175-210.

Murray, Joseph, Rolf Loeber, and Dustin Pardini. 2012. "Parental Involvement in the Criminal Justice System and the Development of Youth Theft Marijuana Use, Depression and Poor Academic Performance." *Criminology*, 50(1): 255-302.

Murray, Joseph, Carl-Gunnar Janson, and David P. Farrington. 2007. "Crime in Adult Offspring of Prisoners: A Cross-National Comparison of Two Longitudinal Samples." *Criminal Justice and Behavior*, 34(1): 133-149.

Norris, S. (2018). *Judicial Errors: Evidence from Refugee Appeals*. University of Chicago, Becker Friedman Institute for Economics Working Paper, (2018-75).

Parke, R. D. and Clarke-Stewart, K. A. (2003). The effects of parental incarceration on children, Prisoners once removed: The impact of incarceration and reentry on children, families, and communities pp. 189–232.

Norris, S., Pecenco, M. and Weaver, J. (2018). "The Effects of Parental and Sibling Incarceration: Evidence from Ohio". Working paper

Pew Center. 2011. "State of Recidivism: The Revolving Door of America's Prisons." Washington, DC: Pew Charitable Trusts.

Kelsey Roberts. 2018. "Fostering Better Educational Outcomes in Youth" Mimeo.

Stith, S.M., Rosen, K.H., Middleton, K.A., Busch, A.L., Lundeberg, K. and Carlton, R.P., 2000. The intergenerational transmission of spouse abuse: A meta-analysis. *Journal of Marriage and Family*, 62(3), pp.640-654.

Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4), 518-529.

Turanovic, J. J., Rodriguez, N., and Pratt, T. C. (2012). The collateral consequences of incar-

ceration revisited: A qualitative analysis of the effects on caregivers of children of incarcerated parents. *Criminology*, 50(4), 913-959.

Vytlacil, E. (2002). Independence, monotonicity, and latent index models: An equivalence result. *Econometrica*, 70(1), 331-341.

Western, Bruce. 2006. *Punishment and Inequality in America*. New York: Russell Sage Foundation.

Western, Bruce and Christopher Muller. 2013. "Mass Incarceration, Macrosociology, and the Poor." *The Annals of the American Academy of Political and Social Science* 647: 166–89.

Western, Bruce, Leonard M. Lopoo, and Sara S. McLanahan. 2004. "Incarceration and the Bonds between Parents in Fragile Families." Pp. 21–45 in *Imprisoning America: The Social Effects of Mass Incarceration*, edited by M. Patillo, D. Weiman, and B. Western. New York: Russell Sage Foundation

Western, B. and Pettit, B.: 2010, *Collateral costs: Incarceration's effect on economic mobility*, Washington, DC: The Pew Charitable Trusts .

Western, B., 2018. *Homeward: Life in the Year After Prison*.

Wildeman, Christopher. 2010. "Paternal Incarceration and Children's Physically Aggressive Behaviors: Evidence from the Fragile Families and Child Wellbeing Study." *Social Forces*, 89(1): 285-309.

Wildeman, Christopher, and Bruce Western. 2010. "Incarceration in Fragile Families." *The Future of Children*, 20(2): 157-177.

Wildeman, Christopher, Signe Hald Anderson, Hedwig Lee and Kristian Bernt Karlson. 2014. "Parental Incarceration and Child Mortality in Denmark." *American Journal of Public Health*, 104(3): 428-433.

2.9 List of Tables

Table 2.1: Population by conviction and incarceration

Sample:	Census: Adult population	SISBEN Criminal record		SISBEN w/ conviction By incarceration	
		No	Yes	No	Yes
	(1)	(2)	(3)	(4)	(5)
Years of education	7.36	6.82	6.68	6.86	6.42
Finished High School D=1	44.0%	31.2%	22.8%	24.2%	20.8%
Income score		34.01	30.90	31.72	29.41
Gender (Male=1)	49.0%	47.6%	83.3%	84.5%	83.3%
# HH members	3.90	4.28	4.47	4.37	4.43
Occupation: Working D=1	48.0%	47.3%	65.4%	67.0%	63.9%
Head of the household D=1		41.2%	47.1%	46.9%	48.6%
Year of birth	1965	1966.9	1974.8	1975.0	1974.3
Marital status: Single D.	45.0%	34.7%	40.7%	45.0%	43.6%
Obs	26,757,687	16,195,178	89,257	55,790	33,467
Years of education for children	8.41	7.20	6.71	6.93	6.57

Notes: Columns 1-5 are group means. HHH: Head of the household, HS: High School. D: Dummy. Income Score: Score from 0 to 100, calculated using variables on income and education of the members of the household, size and characteristics of the house. Source: 2005 Census, SISBEN and criminal records.

Table 2.2: Convicted parents by incarceration and gender

Convicted sample: by gender and incarceration status	Women		Men	
	No	Yes	No	Yes
	(1)	(2)	(3)	(4)
Years of education	6.50	6.06	6.68	6.23
Dummy Has HS degree =1	20%	16%	22%	19%
Income Score	17.2	16.1	19.48	18.46
Occupation: Dummy Working=1	45%	40%	69%	68%
Dummy head of the household=1	36.2%	37.1%	47%	50%
Age at sentence	35.5	36.2	34.46	36.31
Marital status: Dummy Single=1	47.8%	45.1%	46%	44%
Obs	9,375	6,028	46,415	27,439
Notes: Columns 1-4 are group means. HHH: Head of the household, HS: High School. D: Dummy. Income Score: Score from 0 to 100, calculated using variables on income and education of the members of the household, size and characteristics of the house. Source: SISBEN and criminal records.				

Table 2.3: First stage - Parents

Dep var: Decision Dummy	(1)	(2)	(3)	(4)
	Conviction	Conviction	Incarceration	Incarceration
Judge Stringency	0.782*** [0.0143]	0.718*** [0.0204]	0.792*** [0.0416]	0.786*** [0.0430]
Controls	X		X	
F stat*	4.4	4.0	3.9	4.2
F critical value	4	4	4	4
Obs	233,050	116,062	91,854	90,774
Judges	392	392	262	262
R-sq	0.013	0.029	0.243	0.242
adj. R-sq	0.013	0.029	0.238	0.237

Controls column 2: Randomization unit FE, Gender and age. Controls column 4: Randomization unit FE, Gender, YOB FE, Sisben score, year of sentence, court-level and year of survey. Standard errors clustered at the randomization unit level. Sources: Attorney General's Office, criminal records and poverty census. F-stat is calculated from a regression on judge dummies.

Table 2.4: Balance test-Trial sample

Dep. Var: Conviction / Incarceration stringency	Judge: Conviction stringency	Judge: Incarceration stringency
Age	0.0000024 [0.0000208]	0.00000914 [0.0000354]
Gender	0.000324 [0.000509]	-0.000291 [0.000753]
Number of charges	0.000867 [0.000835]	0.000718 [0.00157]
Violent crime	-0.000293 [0.000805]	0.0014 [0.00129]
Property crime	0.00203 [0.00224]	0.00117 [0.00360]
Drugs related crime	-0.000927 [0.00157]	-0.00189 [0.00271]
Guns related crime	-0.000666 [0.00142]	-0.00101 [0.00213]
Misdemeanor	-0.000867 [0.00112]	0.00139 [0.00183]
Obs	187,231	162,960
Judges	1,272	683
F-test	0.52	0.80

Standard errors clustered at the randomization unit/year level. Each rows corresponds to a different regression of judge leniency and defendant characteristics. When testing balance across crime categories I construct an alternative measure of conviction stringency that doesn't parse-out crime level conviction rates. The F-test corresponds to a regression where I include all the variables at the same time. Source Attorney General's office and criminal records.

Table 2.5: Balance test II-Incarcerated sample

Dep var: Incarceration FE	(1) 0.74<Pc<0.88	(2) 0.88<Pc<0.9	(3) 0.9<Pc<1	(4) Pooled Pc
Years of education	-0.0000292 [0.000119]	-0.0000215 [0.000136]	0.000274 [0.000169]	0.00011 [0.0000873]
Income score	-0.0000174 [0.0000283]	0.00000267 [0.0000292]	0.000013 [0.0000364]	0.0000106 [0.0000175]
Age at sentence	0.0000218 [0.0000338]	-2.08E-08 [0.0000320]	0.0000107 [0.0000435]	0.0000197 [0.0000266]
Gender	-0.00142 [0.00127]	0.001 [0.000793]	-0.00212** [0.00100]	-0.00104 [0.000633]
Years of education HH	-0.0000463 [0.000157]	0.000106 [0.000136]	-0.000153 [0.000162]	-0.0000165 [0.0000996]
D: Working	-0.0000919 [0.000672]	-0.000981 [0.000763]	0.000137 [0.00108]	-0.000126 [0.000493]
D: Studying	-0.0022 [0.00316]	-0.000602 [0.00278]	0.00103 [0.00364]	0.00108 [0.00199]
D: Both census surveys	-0.000844 [0.000897]	-0.000942 [0.000634]	0.000587 [0.000857]	-0.000305 [0.000488]
D: First survey	0.000355 [0.00124]	0.000691 [0.00123]	0.000648 [0.00162]	0.000511 [0.000800]
Constant	0.178* [0.107]	-3.04E-01 [0.226]	6.64E-02 [0.124]	0.360*** [0.00594]
F Test	0.8494	0.5001	0.564	0.5763
Obs	16,684	17,416	15,845	49,945
R-sq	0.128	0.149	0.137	0.03
Additional controls: Pc, Randomization unit FE, sentence year FE. Standard errors clustered at the randomization unit year level.				

Table 2.6: Monotonicity test: Norris

Pairwise Monotonicity Test	P-value
Gender	0.33
Primary school	0.99
Young	0.93
Single	0.99
Poor	0.99
Working	0.86
Type of crime:	
Violent	0.52
Property	0.00
Gun-related	0.38
Drug-related	0.32

Norris (2019) test for monotonicity.

Table 2.7: Monotonicity Test: Frandsen et al

Randomization Unit	Critical value	P-value
1	28.561	0.435
2	36.685	0.302
3	22.108	0.279
4	11.612	0.071
5	0.698	0.983
6	5.372	0.372
7	1.197	0.754
8	10.637	0.014
9	2.362	0.501
10	4.485	0.214
11	0.465	0.495
12	0.997	0.607
13	0.265	0.876
14	0.007	0.931
15	4.083	0.130
16	3.72	0.156
Joint test	133.254	0.160

Frandsen et al (2019) test for Monotonicity. I run the test in the randomization units where there are more than 4 judges which corresponds to 73% of my sample.

Table 2.8: OLS Regression

Children with a convicted parent by age 14				
OLS: no controls	(1)	(2)	(3)	(4)
Dep var: Years of education	Pooled Pc	0.7<Pc<0.88	0.88<Pc<0.9	0.9<Pc<1
Parental Incarceration Dummy	-0.356*** [0.0717]	-0.397*** [0.0896]	-0.319*** [0.0815]	-0.372*** [0.0894]
Constant	6.768*** [0.0976]	6.446*** [0.0857]	6.586*** [0.0806]	6.693*** [0.101]
OLS: Adding controls				
Parental Incarceration Dummy	-0.0764*** [0.0235]	-0.0529 [0.0426]	-0.104** [0.0470]	-0.0608* [0.0346]
Obs	52,275	16,091	16,981	16,424
Clusters	609	329	365	403
R-sq	0.383	0.406	0.375	0.364
Controls: Randomization unit FE, Gender, YOB FE, Sisben score, gender of incarcerated parent, pc, year of sentence, birth order and year of survey. Sample: Children between 1990 and 2007 who had a convicted parent between ages 0 and 14. SE in brackets clustered at the randomization unit. AR confidence interval result in the same significance levels.				

Table 2.9: Results: Reduced form and IV

Reduced form	(1)	(2)	(3)	(4)
Dep var: Years of education	Pooled Pc	0.7<Pc<0.88	0.88<Pc<0.9	0.9<Pc<1
Judge leave-out incarceration rate	0.763*** [0.215]	0.720* [0.419]	0.901* [0.503]	0.741** [0.341]
R-sq	0.298	0.322	0.286	0.292
IV Dep var: Years of education	Pooled Pc	0.7<Pc<0.88	0.88<Pc<0.9	0.9<Pc<1
Parental Incarceration Dummy	0.670*** [0.194]	0.670* [0.401]	0.940* [0.552]	0.633** [0.286]
Obs	51,742	16,086	16,979	16,416
Clusters	603	324	363	395

Controls: Randomization unit FE, Gender, YOB FE, Sisben score, gender of incarcerated parent, pc, year of sentence, birth order and year of survey. Column 1 controls add a second order polynomial on Pc. Sample: Children between 1990 and 2007 who had a convicted parent between ages 0 and 14. SE in brackets clustered at the randomization unit. AR confidence interval result in the same significance levels.

Table 2.10: Heterogeneous effects

IV	Girls	Boys	Mother	Father
Dep var: Years of education	(1)	(2)	(3)	(4)
Parental Inc.	0.359* [0.208]	0.865*** [0.286]	0.823** [0.370]	0.531** [0.222]
Obs	26310	27086	12049	41319
	Child's age			
	0-5 years	5-10 years	10-15 years	
Parental Inc.	0.673*** [0.258]	0.384 [0.330]	1.336*** [0.502]	
Obs	18,726	23,634	9,128	
	Type of crime			
	Violent	Property	Drug-related	Gun-related
Parental Inc.	1.634 [1.167]	0.485 [0.606]	0.734** [0.361]	0.542 [0.508]
Obs	9,792	12,985	12,905	9,857
Pooled Pc	x	x	x	x

Controls: Randomization unit FE, Gender, YOB FE, Sisben score, gender of incarcerated parent, pc, year of sentence, birth order and year of survey. Sample: Children between 1990 and 2007 who had a convicted parent between ages 0 and 14. SE in brackets clustered at the randomization unit.

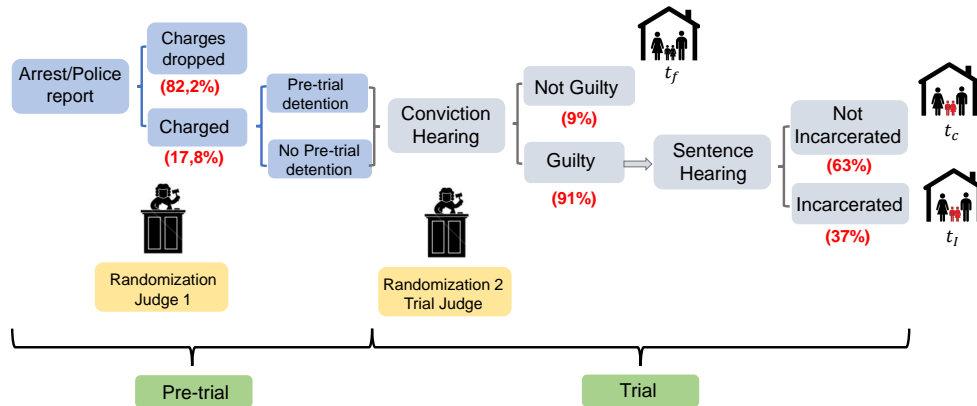
Table 2.11: Changes after incarceration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep var:	LFP spouse	Income score	Years of educ. HHH	D: Male HHH	# of people in HH	D: Lives w/ Grandparents	D: In 2nd SISBEN
Parental Inc.	0.0680*** [0.0187]	-2.365*** [0.193]	0.103*** [0.0300]	-0.0786*** [0.00604]	-0.0996*** [0.0303]	0.0196* [0.0110]	-0.0303*** [0.00492]
Obs	9,673	82,779	82,779	82,779	81,615	16,372	32,881
R-sq	0.22	0.75	0.20	0.19	0.33	0.10	0.08
Mean dep var:	0.399	26.41	5.1	0.595	4.659	0.215	0.242
St dev dep var:	0.49	20.13	2.911	0.491	2.42	0.411	0.428

Panel regressions. Controls: Poverty score, years of education of HHH, Municipality FE and year of survey FE. Dummy for living with grandparents also includes uncles and cousins. Households with data on both cross-sections of the poverty census and who had a conviction episode in between surveys. Source: SISBEN and criminal records.

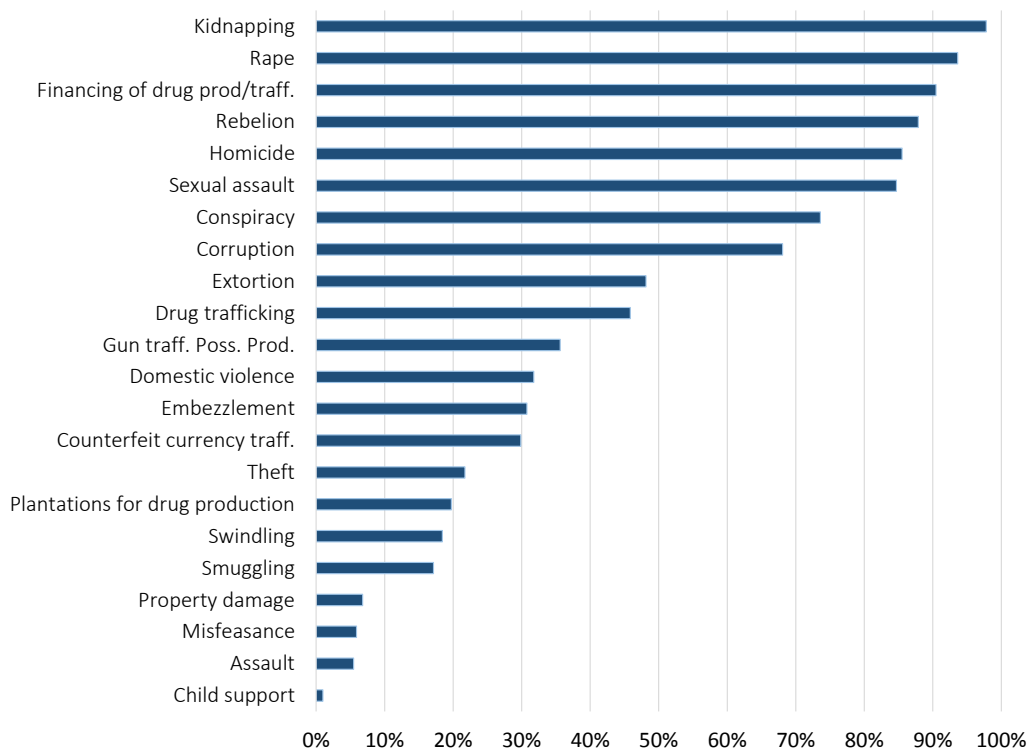
2.10 List of Figures

Figure 2.1: Prosecution and trial stages



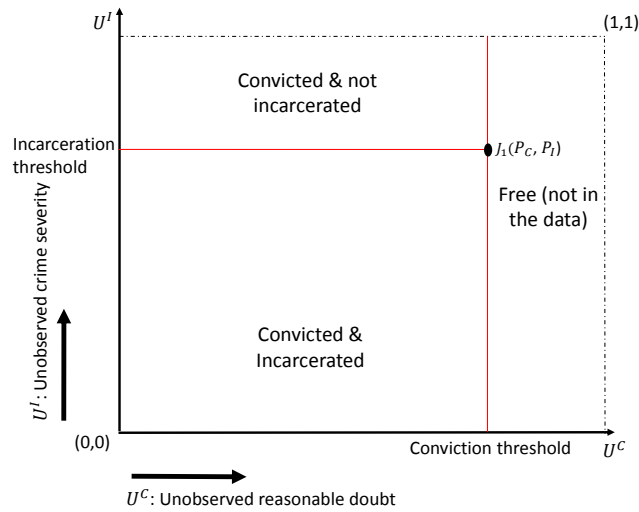
Source: Colombian Penal proceedings code, Informe de la Comision Asesora de Politica Criminal (2012), SPOA and Criminal records. The treatment status studied in this paper corresponds to t_f , which refers to parents who are not convicted or free, t_c those convicted but not incarcerated, and t_l those convicted and incarcerated. Incarceration is a function of sentence length. Currently, a sentence equal to four years or less is not served in prison.

Figure 2.2: Incarceration rates



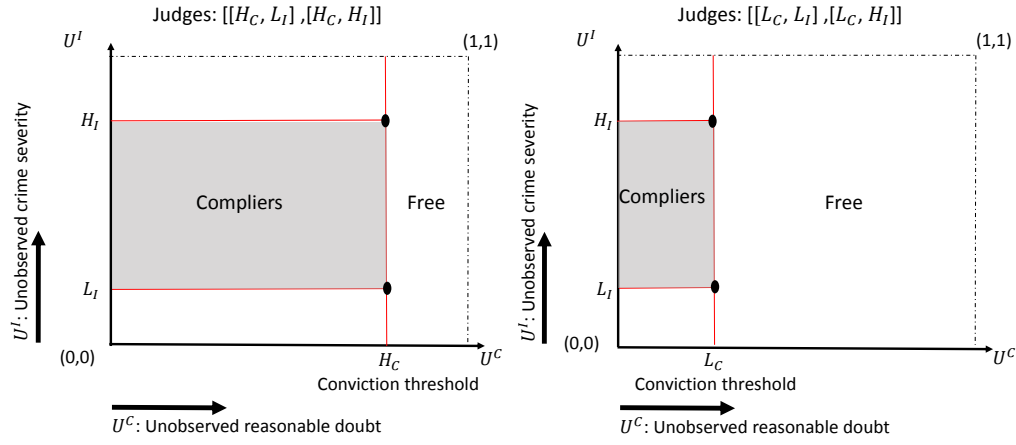
Source: Criminal records. Selected crimes. I restrict to crimes with at least 100 cases.

Figure 2.3: Identification: Defendant's space, judges thresholds and treatment assignment



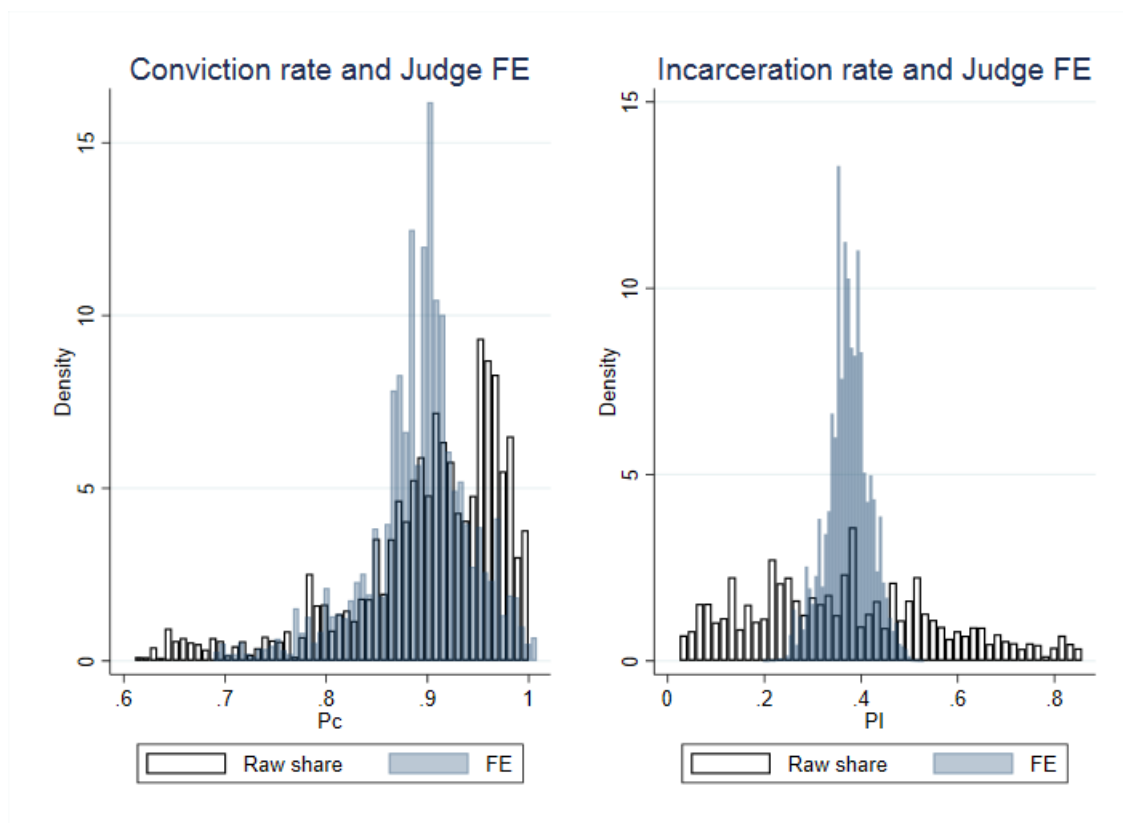
A defendant is characterized by a point in the unitary square. A judge is defined by a pair of thresholds along each axis which determine treatment assignments. Defendants to the left of the conviction threshold are convicted, and those to the right are freed. Among the convicted, defendants below the incarceration threshold go to prison, and those above do not.

Figure 2.4: Identification under 4 types of judges



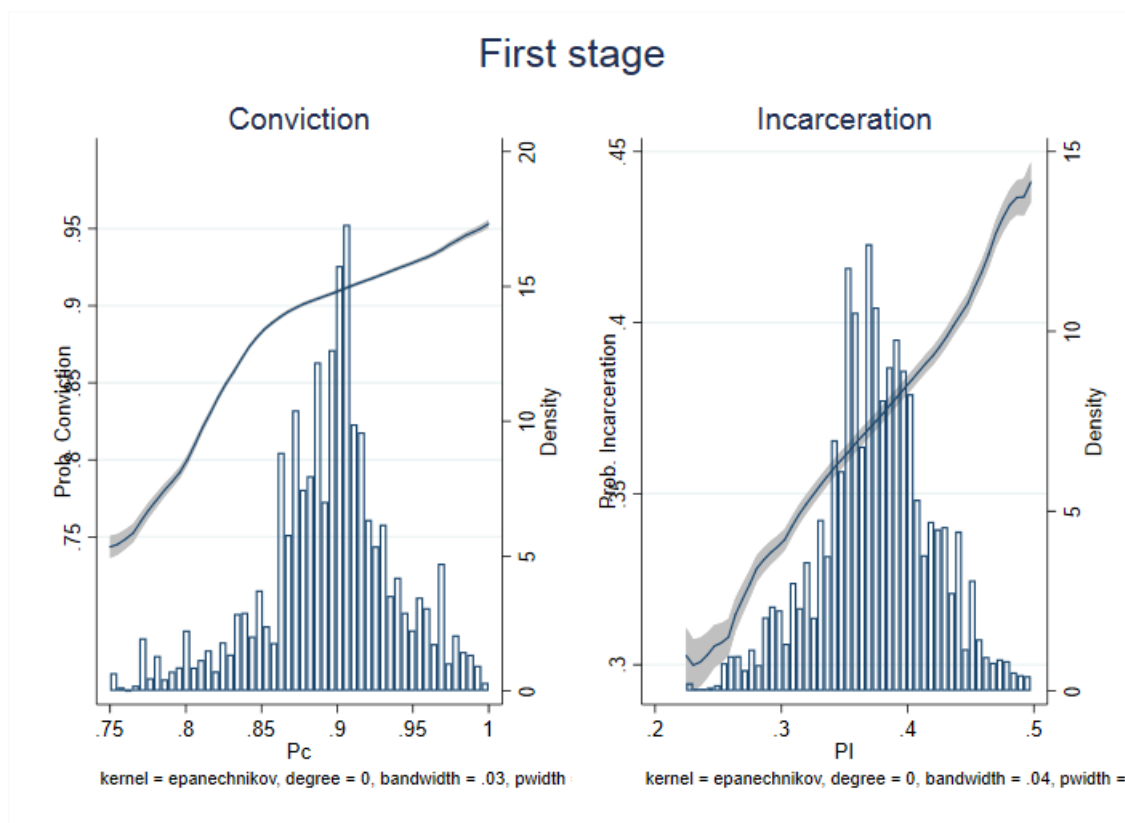
The left panel features harsh judges on the conviction margin (H_C). This judges can be harsh (H_I) or lenient (L_I) on the incarceration margin. We can identify the causal effect of incarceration for defendants in the shaded area. Those whose incarceration decision is only a function of judge assignment. The right panel is analogous and it features lenient judges on the conviction margin (L_C).

Figure 2.5: Judges' fixed effects



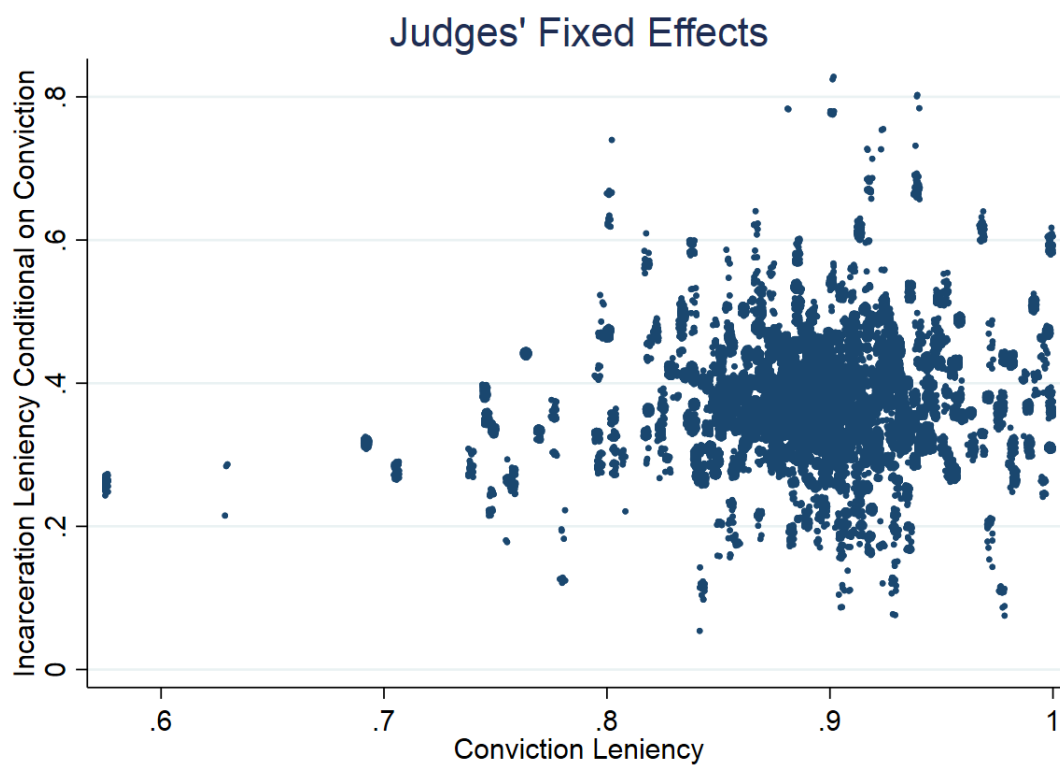
Source: Attorney General's office and criminal records. Raw rates are conviction/incarceration averages by judge. To construct the judge's fixed effect I take the residual at the judge level after regressing conviction/incarceration on (demeaned) randomization unit/year dummies, (demeaned) crime-level conviction/incarceration rates, without a constant.

Figure 2.6: First stage



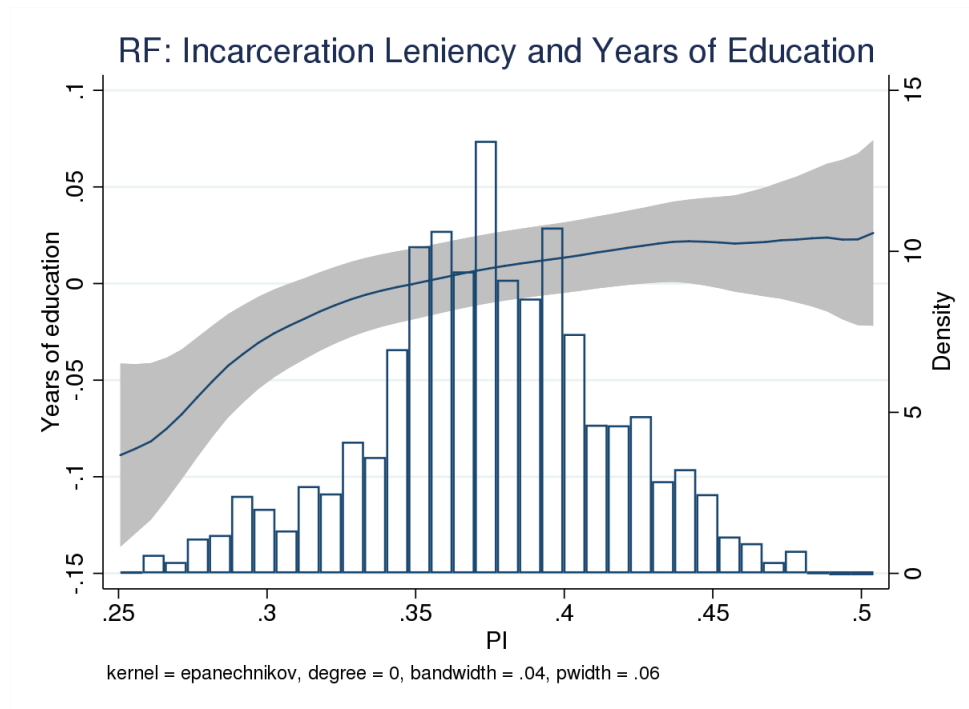
Source: Attorney General's office and criminal records. Raw rates are conviction/incarceration averages by judge. To construct the judge's fixed effect I take the residual at the judge level after regressing conviction/incarceration on (demeaned) randomization unit/year dummies, (demeaned) crime-level conviction/incarceration rates, without a constant.

Figure 2.7: Scatter plot: Judges' fixed effects



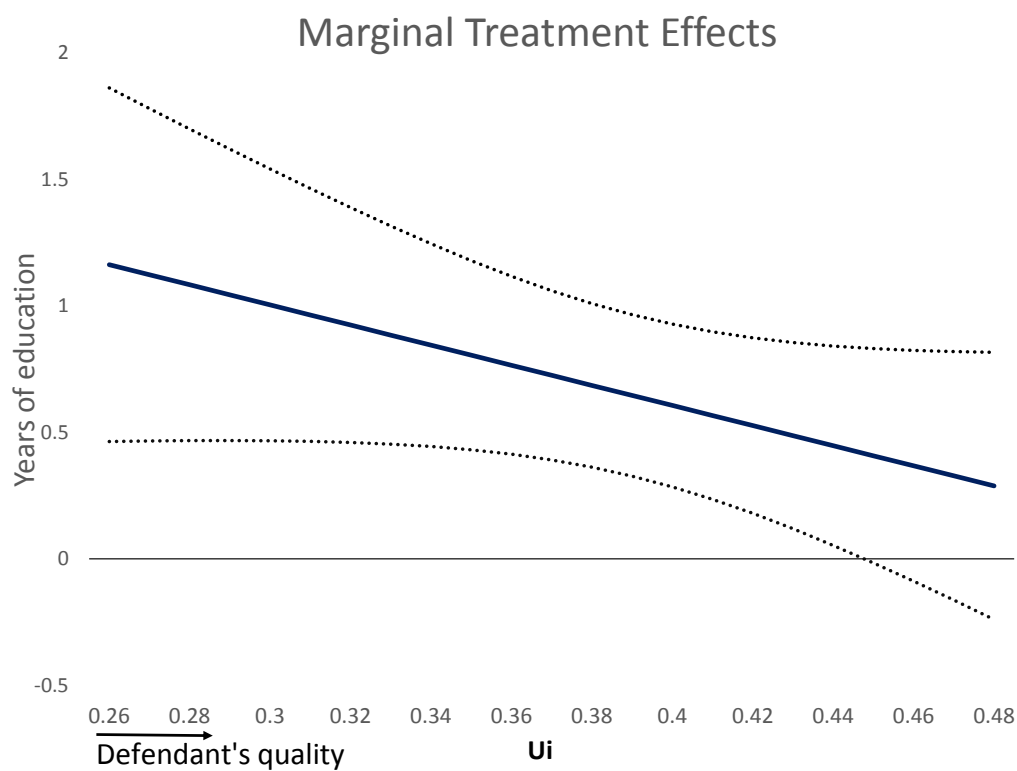
Source: Attorney General's office and criminal records. To construct the judge's fixed effect I take the residual at the judge level after regressing conviction/incarceration on (demeaned) randomization unit/year dummies, (demeaned) crime-level conviction/incarceration rates, without a constant.

Figure 2.8: Reduced form



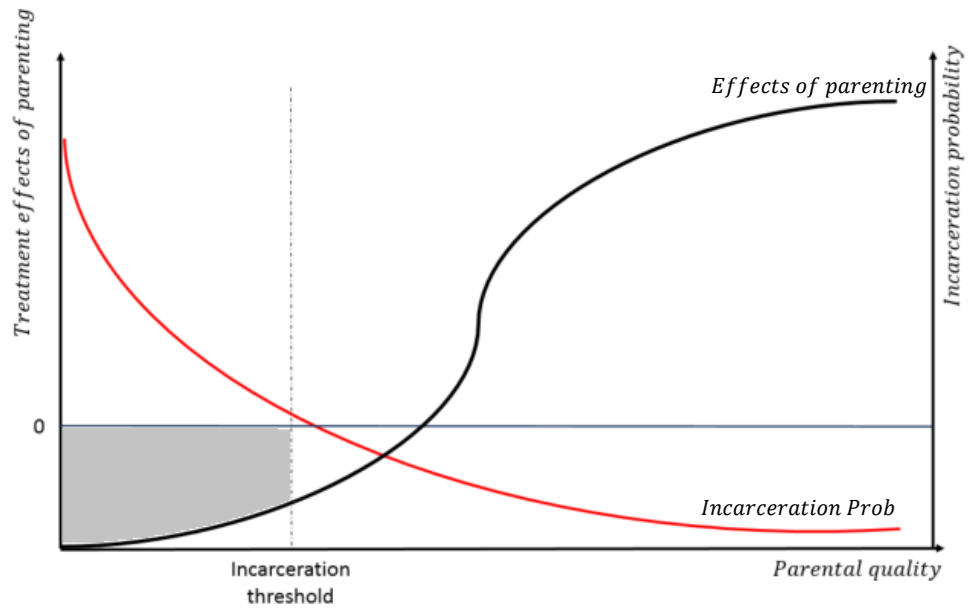
Notes: Histogram of parental incarceration judge leniency and the fitted value of local polynomial regressions of children's educational attainment on judge stringency. Pooled regression I control for p_c .

Figure 2.9: MTE



Notes: Following the LIV approach in Heckman and Vytlačil (2005) I regress $Yeduc = \alpha + \beta_1 P_i + \beta_2 P_i^2 + \beta_3 X$. This graphs plots: $\beta_1 + 2\beta_2 P_i$ for the pooled regression. Controls: Municipality FE, gender, YOB FE, Sisben score, years of education HH head, years of education of incarcerated parent, gender of incarcerated parent, pc, year of sentence, birth order and year of survey.

Figure 2.10: Model of parenting and incarceration



Chapter 3

Identification of treatment effects

3.1 Model

In this chapter, I formalize the previous intuition and extend it to the case of continuous instruments to deliver a new identification result.

The model is described by the standard IV model, which consists of five main random variables: $T, Z, Y, \mathbf{V}, \mathbf{X}$. Those variables lie in the probability space (Ω, F, P) , where individuals are represented by elements $i \in \Omega$ of the sample space Ω . The variables are defined below:

- T_i denotes the assigned treatment of individual i , and takes values in $\text{supp}(T) = \{t_f, t_c, t_l\}$. t_f stands for not convicted, t_c for convicted but not incarcerated, and t_l for convicted and incarcerated.
- Z_i is the instrumental variable in this analysis and takes values in $\text{supp}(Z)$, and represents judge assignment.
- Y_i denotes the outcome of interest for individual i , —e.g., years of education of the child.
- \mathbf{X}_i represents the exogenous characteristics of individual i .
- \mathbf{V}_i stands for the random vector of unobserved characteristics of individual i , and takes values in $\text{supp}(\mathbf{V})$.

The random vector \mathbf{V} is the source of selection bias in this model. It causes both the treatment T and outcome Y . The standard IV model is defined by two functions and an independence condition, as follows:

$$\text{Outcome Equation: } Y = f_Y(T, \mathbf{X}, \mathbf{V}, \varepsilon_Y) \quad (3.1)$$

$$\text{Treatment Equation: } T = f_T(Z, \mathbf{X}, \mathbf{V}) \quad (3.2)$$

$$\text{Independence: } Z \perp (\mathbf{V}, \varepsilon_Y) | \mathbf{X} \quad (3.3)$$

where ε_Y is an unobserved zero-mean error term associated with the outcome equation.

In this notation, a counterfactual outcome is defined by fixing T to a value $t \in \text{supp}(T)$ in the outcome equation. That is, $Y(t) = f_Y(t, \mathbf{V}, \mathbf{X}, \varepsilon_Y)$. The observed outcome for individual i is given by:

$$Y = Y(T) = \sum_{t \in \{t_f, t_c, t_l\}} Y(t) \cdot \mathbf{1}[T = t]. \quad (3.4)$$

The independence condition (3.3) implies the following exclusion restriction:

$$\text{Exclusion Restriction : } Z \perp Y(t) | \mathbf{X} \text{ for all } t \in \text{supp}(T). \quad (3.5)$$

For the sake of notational simplicity, I suppress exogenous variables \mathbf{X} henceforth. All of the analysis can be understood as conditional on pre-treatment variables.

I assume that the treatment equation is governed by a combination of two threshold-crossing inequalities. First, there is a conviction stage:

$$\begin{cases} \text{Free} & \text{if } \mathbf{1}[\phi_c(\mathbf{V}) > \xi_c(Z)] \\ \text{Convicted} & \text{if } \mathbf{1}[\phi_c(\mathbf{V}) \leq \xi_c(Z)]. \end{cases}$$

where $\mathbf{1}[\cdot]$ denotes a binary indicator and $\phi_c(\cdot), \xi_c(\cdot)$ are real-valued functions. Function $\phi_c(\cdot)$ measures the degree of culpability assessed by the judicial system. This function looks at variables and information that are not observed by the econometrician but that are observed by the judge, such as the evidence, crime intensity, effort of the defense and prosecutor lawyers, as well as unobserved characteristics of the defendant such as aggression, antisocial behavior, etc. The function $\xi_c(\cdot)$ assesses judge leniency on conviction. This function can be understood as a threshold of reasonable doubt beyond which the defendant is convicted by the judge. Judges differ in their leniency and may set different thresholds for evidence. The judge convicts defendant i whenever $\phi_c(V_i) \leq \xi_c(Z_j)$. If that is the case, a second stage is held and the judge makes a decision regarding incarceration:

$$\begin{cases} \text{Not incarcerated} & \text{if } \mathbf{1}[\phi_I(\mathbf{V}) > \xi_I(Z)] \\ \text{Incarcerated} & \text{if } \mathbf{1}[\phi_I(\mathbf{V}) \leq \xi_I(Z)] \end{cases}$$

Similarly, $\phi_I(\mathbf{V})$ is a function whose arguments are the case and defendant's characteristics relevant for assessment of the punishment level. As before, the judge compares $\phi_I(\mathbf{V})$ to her/his threshold to incarcerate $\xi_I(Z)$.

Treatment assignment can be summarized as follows:¹

$$T = f_T(Z, \mathbf{V}) = \begin{cases} t_f & \text{if } \mathbf{1}[\phi_c(\mathbf{V}) > \xi_c(Z)] \\ t_c & \text{if } \mathbf{1}[\phi_c(\mathbf{V}) \leq \xi_c(Z)] \cdot \mathbf{1}[\phi_I(\mathbf{V}) > \xi_I(Z)] \\ t_I & \text{if } \mathbf{1}[\phi_c(\mathbf{V}) \leq \xi_c(Z)] \cdot \mathbf{1}[\phi_I(\mathbf{V}) \leq \xi_I(Z)] \end{cases}$$

¹I assume the following standard regularity conditions: A1) $E(|Y(t)|) < \infty$ for all $t \in \text{supp}(T)$, A2) $P(T = t|Z = z) > 0$ for all $t \in \text{supp}(T)$ and all $z \in \text{supp}(Z)$ and, A3) $(\phi_c(\mathbf{V}), \phi_I(\mathbf{V}))$ are absolutely continuous with respect to Lebesgue measure in \mathbb{R}^2 . The first assumption guarantees the existence of the expectation, the second one assures that there is a share of the population assigned to each treatment group for every judge, and the third one allows me to apply the Lebesgue differentiation theorem.

This model relies on two separable threshold functions that play the role of the monotonicity condition (Vytlačil, 2002).² Without loss of generality, it is useful to express treatment assignment using the following variable transformation:

$$U^c = F_{\phi^c(\mathbf{V})}(\phi^c(\mathbf{V})) \sim Unif[0, 1], \quad (3.6)$$

$$U^I = F_{\phi^I(\mathbf{V})}(\phi^I(\mathbf{V})) \sim Unif[0, 1] \quad (3.7)$$

where $F_K(\cdot)$ denotes the cumulative distribution function of a random variable K . U^c, U^I are uniformly distributed random variables in $[0, 1]$ due to assumption A3, and there is no restriction on the joint distribution of U^I and U^c .

$$P_c(z) = F_{\phi^c(\mathbf{V})}(\xi^c(Z)); z \in \text{supp}(Z), \quad (3.8)$$

$$P_I(z) = F_{\phi^I(\mathbf{V})}(\xi^I(Z)); z \in \text{supp}(Z) \quad (3.9)$$

Let $P_c(z)$ denote the probability of conviction when $Z = z$. Moreover, independence condition (3.3)

²Consider two judges, j and j' , who see defendants i and i' , who differ in their level of culpability. Say i' has more evidence against him than i ; namely $\phi_c(i') < \phi_c(i)$. Suppose that judge j convicts defendant i' but not i . Then the threshold function implies that it cannot be the case that judge j' convicts defendant i , but not i' . More generally, let $D_i(j) = \mathbf{1}[T_i(j) = t_c]$ denote the binary indicator that judge j convicts defendant i . Thus if judge j convicts i' but not i , it implies:

$$D_i(j) > D_{i'}(j)$$

Then it cannot be the case that judge j' convicts defendant i , but not i' . Which means:

$$D_i(j) > D_{i'}(j) \rightarrow D_i(j') \geq D_{i'}(j'),$$

which is equivalent to stating that:

$$D_i(j) > D_i(j') \rightarrow D_{i'}(j) \geq D_{i'}(j').$$

We can generalize this to all individuals to arrive at the standard monotonicity assumption of Imbens and Angrist (1994).

implies $P_c, P_I \perp U^c, U^I$. In this notation, the model can be expressed as:

$$T \equiv f_T(Z, \mathbf{V}) = g_T(U^c, U^I, P_c, P_I) = \begin{cases} t_f & \text{if } \mathbf{1}[U^c > P_c(z)] \\ t_c & \text{if } \mathbf{1}[U^c \leq P_c(z)] \cdot \mathbf{1}[U^I > P_I(z)] \\ t_I & \text{if } \mathbf{1}[U^c \leq P_c(z)] \cdot \mathbf{1}[U^I \leq P_I(z)] \end{cases} \quad (3.10)$$

For ease of exposition, I will first explore identification under the assumption that $U^c \perp U^I$ and then I will go over the results without it. Under the independence assumption we can identify $P_I(z)$ from the data, that is:

$$P(U^I < P_I(z) | U^c \leq P_c(z)) = P(U^I < P_I(z)) = P_I(Z)$$

The left hand side is observed from the data, the first equality follows directly from the independence assumption and the last one the uniform distribution of U^I . P_I is interpreted as the share incarcerated.

The goal is to identify and evaluate the treatment effect: $E(Y(t_I) - Y(t_c))$ which is a function of counterfactual variables $Y(t_I)$ and $Y(t_c)$. To achieve this goal, it is useful to express the observed expectations in terms of the variables that define the model:

$$E(Y \cdot \mathbf{1}[T = t_c] | P_c(Z) = p_c, P_I(Z) = p_I) = \quad (3.11)$$

$$= E(Y(t_c) \cdot \mathbf{1}[T = t_c] | P_c(Z) = p_c, P_I(Z) = p_I) \quad (3.12)$$

$$= E(Y(t_c) \cdot \mathbf{1}[U^c \leq p_c] \cdot \mathbf{1}[U^I > p_I] | P_c(Z) = p_c, P_I(Z) = p_I) \quad (3.13)$$

$$= E(Y(t_c) \cdot \mathbf{1}[U^c \leq p_c] \cdot \mathbf{1}[U^I > p_I]) \quad (3.14)$$

$$= \int_0^{p_c} \int_{p_I}^1 E(Y(t_c) | U^c = u^c, U^I = u^I) f_{u^c u^I}(u^c, u^I) du^c du^I \quad (3.15)$$

$$(3.16)$$

$$= - \int_0^{p_c} \int_0^{p_I} E(Y(t_c)|U^c = u^c, U^I = u^I) f_{u^c, u^I}(u^c, u^I) du^c du^I + \int_0^{p_c} E(Y(t_c)|U^c = u^c) f_{u^c}(u^c) du^c$$

Equation (3.12) is an expectation observed in the data. Equality (3.13) comes from the definition of observed outcomes. Equality (3.14) expresses the indicator $\mathbf{1}[T = t_c]$ in terms of the inequalities of the choice model. Equality (3.15) uses the independence relation $Z \perp (U^c, U^I)$. Equality (3.16) expresses the expectation as the integral over the distribution of U^c, U^I where $f_{U^c, U^I}(u^c, u^I)$ stands for the probability density function of U^c, U^I at the point (u^c, u^I) , and is equal to one. Equality (3.19) modifies the integration region. This change is useful to apply the Lebesgue differentiation theorem next;

$$\frac{\partial^2 E(Y \cdot \mathbf{1}[T = t_c] | P_c(Z) = p_c, P_I(Z) = p_I)}{\partial p_c \partial p_I} = -E(Y(t_c) | U^c = p_c, U^I = p_I) \quad (3.17)$$

Equality (3.17) arises as a direct application of the Lebesgue differentiation theorem. What this result gives me is a connection between the observed outcomes and the targeted counterfactual outcome. We can use the same steps applied to counterfactual $Y(t_c)$ to obtain the counterfactual for $Y(t_I)$. Combining these two I obtain:

$$\frac{\partial^2 E(Y \cdot \mathbf{1}[T \in \{t_c, t_I\}] | P_c(Z) = p_c, P_I(Z) = p_I)}{\partial p_c \partial p_I} = E(Y(t_I) - Y(t_c) | U^c = p_c, U^I = p_I) \quad (3.18)$$

In the language of Heckman and Vytlacil (2005), Eq.3.18 defines the marginal treatment effect (MTE) of outcome Y with respect to treatment assignment t_c and t_I . It is interpreted as the causal effect of incarceration versus conviction only, for the share of defendants whose culpability and punishment assessments, U^c and U^I respectively, is set at quantiles p_c and p_I . The derivative in Equation (3.16) traces the MTE of incarceration relative to conviction throughout the unitary square of U^c, U^I . This result is an application of Lee and Salanie (2018) and extends the result

of Heckman and Vytlacil (1999). In Appendix B I explain graphically the intuition of this result. The main idea is that changes in P_c and P_I affect exogenously treatment assignment. Then, by examining the derivative of the outcome variables with respect to P_c and P_I , we capture how the outcome variable changes when treatment changes at each point in the space of the unobservable confounding variables.

The average treatment effect (ATE) is the causal effect of t_c and t_I on Y in the population, and it corresponds to the integral of the MTE over the support of U^c and U^I .

$$E(Y(t_I) - Y(t_c)) = \int_0^1 \int_0^1 \frac{\partial^2 E(Y \cdot \mathbf{1}[T \in \{t_c, t_I\}] | P_c(Z) = p_c, P_I(Z) = p_I)}{\partial p_c \partial p_I} dp_c dp_I \quad (3.19)$$

Without the assumption of independence of U^c and U^I , variation in P_I is only identified once I fix the conviction threshold. Thus, the counterfactual of interest is now: $Y(t_I)$ and $Y(t_c)$ for those who were convicted under $P_c = p_c$. This means the objective is to identify causal effects of the form: $E(Y(t_I) - Y(t_c) | U^c < p_c)$, which is the the same exercise explained in Section 4.1. Let:

$$E(Y \cdot \mathbf{1}[T = t_c] | P_c(Z) = p_c, P_I(Z) = p_I, U^c < p_c) = \quad (3.20)$$

$$= E(Y(t_c) \cdot \mathbf{1}[T = t_c] | P_c(Z) = p_c, P_I(Z) = p_I, U^c < p_c) \quad (3.21)$$

$$= E(Y(t_c) \cdot \mathbf{1}[U^I > p_I] | P_c(Z) = p_c, P_I(Z) = p_I, U^c < p_c) \quad (3.22)$$

$$= E(Y(t_c) \cdot \mathbf{1}[U^I > p_I] | U^c < p_c) \quad (3.23)$$

Where I followed the same steps as before. Let:

$$P_I^* = Pr[U^I < P_I | U^c < P_c] = G(P_I) \quad (3.24)$$

P_I^* is the object I observe so I will define the observed expectations in terms of this variable:³

$$E(Y(t_c) \cdot \mathbf{1}[U^I > G^{-1}(p_I^*|U^c < p_c)]|U^c < p_c) \quad (3.25)$$

$$\int_{P_I^*}^1 E(Y(t_c)|U^I = u^I, U^c < p_c) f_{u^I|U^c < p_c}(p_I^*) du^I \quad (3.26)$$

And applying the Lebesgue differentiation theorem this results in:

$$\frac{\partial E(Y \cdot \mathbf{1}[T \in \{t_c\}]|p_c, p_I, U^c < p_c)}{\partial p_I^*} = -E(Y(t_c)|U^I = p_I, U^c < p_c) f_{u^I|U^c < p_c}(p_I^*) \quad (3.27)$$

And ultimately;

$$E(Y(t_I) - Y(t_c)|U^c < p_c) = \int_0^1 \frac{\partial E(Y \cdot \mathbf{1}[T \in \{t_c, t_I\}]|P_c(Z) = p_c, P_I^*(Z) = p_I^*, U^c < p_c)}{\partial p_I^*} dp_I^* \quad (3.28)$$

What this result says is that we can trace the treatment effect of incarceration relative to conviction once we fix a threshold for conviction. We do this by evaluating the changes on the outcome variable when we change P_I^* . This delivers the MTE along the unobservable dimension $U^I|U^c < P_c$. The integral over the support of the instrument gives the LATE, or the ATE when the instrument has full support.

3.2 References

Ahn, H., & Powell, J. L. (1993). Semiparametric estimation of censored selection models with a nonparametric selection mechanism. *Journal of Econometrics*, 58(1-2), 3-29.

Fomin, S. V. (1999). *Elements of the theory of functions and functional analysis* (Vol. 1). Courier Corporation.

Imbens, G.W., and D. B. Rubin (1997). Estimating Outcome Distributions for Compliers in Instrumental Variables Models,” *The Review of Economic Studies*, 64(4).

³Where $f_{u^I|U^c < p_c}(p_I^*)$ in eq. (39) corresponds to: $f_{u^I|U^c < p_c}(p_I) \frac{\partial P_I((p_I^*))}{(p_I^*)}$

Heckman, J. J., Urzua, S., and Vytlačil, E. (2006). Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics*, 88(3), 389-432.

Heckman, James J., and Edward Vytlačil. 2005. "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica*, 73(3): 669-738.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Q. Yavitz (2010a). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative Economics* 1 (1), 1–46.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Q. Yavitz (2010b). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics* 94 (1–2), 114–128.

Heckman, J. J. (2013). *Giving kids a fair chance*. Mit Press.

Lee, Sokbae, and Bernard Salanié. "Identifying effects of multivalued treatments." (2018).